DOCUMENT RESUME

ED 166 638 CS 004 563

AUTHOR	Sternberg, Robert J.
TITLE	Representation and Process in Transitive Inference. Technical Report No. 12.
INSTITUTION	Yale Univ., New Haven, Conn. Dept. of Psychology.
SPONS AGENCY	National Inst. of Mental Health (DHEW), Rockville, Md.; National Science Foundation, Washington, D.C.; Office of Naval Research, Arlington, Va. Personnel and Training Research Programs Office.
PUB DATE	Oct 78
CONTRACT	N0001478C0025
GRANT	BNS-76-05311; MH-13950-07
NOTE	125p.; Several figures may not reproduce due to small type
EDRS PRICE	MF-$0.83 HC-$6.01 Plus Postage.
DESCRIPTORS	Abstraction Levels; *Association (Psychological); *Cognitive Processes; Cognitive Style; Cues; Information Processing; Learning Processes; *Linguistics; *Memory; Psychological Studies; Retention; Thought Processes
IDENTIFIERS	*Transitive Inference

ABSTRACT

In this report, three theories of transitive inference are compared as they apply to the solution of linear syllogisms: a spatial theory, a linguistic theory, and a new mixed linguistic-spatial theory. Each theory is expressed in terms of an information-processing (flow chart) model and a mathematical model that quantifies the information-processing model. The results of four experiments testing the mathematical models for their ability to account for latency data are reported, and it is shown that they overwhelmingly support the mixed theory. This support holds over varied modes of problem presentation, adjectives, sessions, and subjects. The duration of each component process in the mixed theory is estimated, and its contribution to total solution time is assessed. Then the mixed theory is shown to account for most patterns of individual-difference data. Finally, the theory is shown to be consistent with a variety of data obtained in previous investigations of transitive inference. (Fourteen tables are appended.) (Author/MAI)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY.

ED166638

Representation and Process in Transitive Inference

Robert J. Sternberg

Yale University

Technical Report No. 12
October, 1978

Approved for public release; distribution unlimited.
Reproduction in whole or in part is permitted for
any purpose of the United States Government.

This research was sponsored by the Personnel and Training Research Programs, Psychological Sciences Division, Office of Naval Research, under Contract No. N0001478C0025, Contract Authority Identification Number NR 150-412.

Send proofs to Robert J. Sternberg
Department of Psychology
Yale University
Box 11A Yale Station
New Haven, Connecticut 06520

Running Head: Transitive Inference

September, 1978

CS004563

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER
Technical Report No. 12

2. GOVT ACCESSION NO.

3. RECIPIENT'S CATALOG NUMBER

4. TITLE (and Subtitle)
"Representation and Process in Transitive Inference"

5. TYPE OF REPORT & PERIOD COVERED
Periodic Technical Report
(1 July 78 - 30 Sept 78)

6. PERFORMING ORG. REPORT NUMBER
Research Report No. 12-78

7. AUTHOR(s)
Robert J. Sternberg

8. CONTRACT OR GRANT NUMBER(s)
N0001478C0025

9. PERFORMING ORGANIZATION NAME AND ADDRESS
Department of Psychology
Yale University
New Haven, Connecticut 06520

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
61153N;
RR 042-04; RR 042-04-01;
NR 150-412

11. CONTROLLING OFFICE NAME AND ADDRESS
Personnel and Training Research Programs
Office of Naval Research (Code 458)
Arlington, Virginia 22217

12. REPORT DATE
30 Oct 78

13. NUMBER OF PAGES
114

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)

15. SECURITY CLASS. (of this report)
Unclassified

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)
Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES
The conduct of this research was supported by NSF Grant BNS-76-05311 (Experiments 2-4) to Robert J. Sternberg and NIMH Grant MH-13950-07 to Gordon H. Bower (Experiment 1). Preparation of this report was supported by ONR Contract N0001478C0025 and NSF Grant BNS-76-05311 to Robert J. Sternberg.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)
transitive inference, linear syllogistic reasoning, three-term series problem, deduction, mixture theory

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)
This article compares three theories of transitive inference applied to the solution of linear syllogisms: a spatial theory, a linguistic theory, and a new mixed linguistic-spatial theory. Each theory is expressed in terms of an information-processing (flow-chart) model, and a mathematical model that quantifies the information-processing model. The mathematical models are tested in their ability to account for latency data from four experiments. The tests overwhelmingly support the mixed theory. This support

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-LF-014-6601

holds over varied modes of problem presentation, adjectives, sessions, and subjects. The duration of each component process in the mixed theory is estimated, and its contribution to total solution time assessed. Then the mixed theory is shown to account for most patterns of individual-difference data. Finally, the theory is shown to be consistent with a variety of data obtained in previous investigations of transitive inference.

Abstract

This article compares three theories of transitive inference applied to the solution of linear syllogisms: a spatial theory, a linguistic theory, and a new mixed linguistic-spatial theory. Each theory is expressed in terms of an information-processing (flow-chart) model, and a mathematical model that quantifies the information-processing model. The mathematical models are tested in their ability to account for latency data from four experiments. The tests overwhelmingly support the mixed theory. This support holds over varied modes of problem presentation, adjectives, sessions, and subjects. The duration of each component process in the mixed theory is estimated, and its contribution to total solution time assessed. Then the mixed theory is shown to account for most patterns of individual-difference data. Finally, the theory is shown to be consistent with a variety of data obtained in previous investigations of transitive inference.

Representation and Process in Transitive Inference

In a transitive inference problem, an individual is presented with two or more premises, each describing a relation between two items. At least one of the items overlaps between premises. The subject's task is to use this overlap to determine a relation between two (or more) items not occurring in the same premise. An example of such a problem is

Mighty Joe Young is mightier than King Kong.

King Kong is mightier than Magilla Gorilla.

Who is mightiest?

This problem illustrates a simple form of transitive inference problem, the linear syllogism (or three-term series problem, as it is often called). In general, the terms of a linear syllogism form a linear array of items, say, ($\underline{A}$, $\underline{B}$, $\underline{C}$). Each of two premises describes a relation between one pair of adjacent items, say, ($\underline{A}$ r_1 $\underline{B}$), ($\underline{B}$ r_2 $\underline{C}$). To solve the problem, an individual must combine information from the two premises in order to determine the relation between the two nonadjacent items, ($\underline{A}$ r_3 $\underline{C}$). Solution of the example problem above requires the individual to infer the relation between the two nonadjacent items, Mighty Joe Young and Magilla Gorilla.

Transitive inferences are widely used in everyday life. Comparisons and decisions of almost every kind that we make on a daily basis usually involve at least an implicit transitive inference. Consider, for example, the plight of a customer eating at a restaurant. The customer is faced with what may well be a bewildering choice of meals. The customer has neither the time nor the patience to compare every possible pair of meals, and to order the meal that is preferred to every other. More typically, the customer will narrow down his or her preferences to a few possible choices, assuming that if the eliminated choices are less desirable than the minimally acceptable choice, they are less desirable than any other acceptable choice as well. Next, the customer will probably pick one of the, say, four semifinal choices as the final choice. Again, the customer eschews making all possible paired comparisons, six in the case of four choices. Instead, the customer assumes transitivity of preferences, and infers that if his or her first choice is preferred to the second choice, it is preferred to every other choice as well. Without the use of transitive inference, many of even our simplest decisions would become unmanageably complex.

Psychologists have long recognized the fundamental importance of transitive inferences in everyday cognition; as a result, transitive inference has played a key role in psychological theory. Research on transitive inference has appeared in diverse psychological literatures, and under a number of different guises. Differential psychologists have recognized the transitive inference problem as a useful psychometric tool since Burt's (1919) use of the problem in a battery of mental tests, although our knowledge of the psychometric properties of the problem as a test item remains rudimentary (Burt, 1919; Shaver, Pierson, & Lang, 1974). Developmental psychologists have investigated the transitive inference problem extensively, many of them in response to Piaget's (1921, 1928, 1955, 1970) claim that preoperational children are unable to perform the reasoning necessary to infer a transitive relation. Trabasso (Bryant & Trabasso, 1971; Riley & Trabasso, 1974), for example, has taken issue with Piaget's interpretation of the data, and in a series of ingenious experiments has found evidence suggesting that memory rather than reasoning limitations are responsible for much of the difficulty young children encounter in attempting to solve transitive inference problems.

Cognitive-experimental psychologists have also devoted a great deal of attention to the transitive inference problem, and have engaged in a vigorous debate regarding the representations and processes subjects use in solving such problems (Clark, 1969a, 1969b, 1971, 1972a, 1972b; DeSoto, London, & Handel, 1965; Handel, DeSoto, & London, 1968; Hunter, 1957; Huttenlocher, 1968; Huttenlocher & Higgins, 1971, 1972; Huttenlocher, Higgins, Milligan, & Kauffman, 1970; Potts & Scholz, 1975). In this article, a new resolution of the debate regarding representation and process will be proposed. In particular, the article will address three important theoretical questions:

1. How is information represented during the course of a subject's solution of a transitive inference problem?

2. What processes act upon the representation(s) from the time the subject first views the problem to the time he or she indicates a response?

3. How, if at all, do the representations and processes used in the solution of transitive inference problems vary as a function of (a) mode of problem presentation, (b) adjectives, (c) sessions, and (d) individual subjects?

The article is divided into six major parts, of which the first is this introduction. The second part presents a discussion of previous research on the three theoretical questions posed above. The third part offers three detailed information-processing models that purport to describe how subjects solve transitive inference problems with three terms (linear syllogisms). The fourth part presents results of four experiments addressed to the theoretical questions posed above, and which are intended to distinguish among the three information-processing models. The fifth part of the article contains tests of the models on previously published data. The sixth and final part of the paper discusses how the data presented in the article advance our theoretical understanding of transitive inference.

THEORETICAL ISSUES

Representation of Information in Transitive Inference

Nature of Controversy

Theorists disagree as to the form of representation subjects use for information stored, manipulated, and retrieved in the course of solving transitive inference problems. The basic controversy has been over whether information is represented spatially or linguistically. Spatial theorists argue that information is represented in the form of a spatial array that functions as an internal analogue to a physically realized or realizable array. Linguistic theorists argue that information is represented in the form of linguistic, deep-structure propositions of the type originally proposed by Chomsky (1965). A resolution of this controversy would not only enlighten us with regard to transitive inference, but might further shed light on the kinds of arguments that are valuable in distinguishing between subjects' use of spatial or imagerial representations on the one hand, and linguistic or propositional representations on the other (see Kosslyn & Pomerantz, 1977; Pylyshyn, 1973; Anderson, Note 1). In the following four sections, evidence for each kind of representation is presented and then evaluated.

Evidence for Spatial Representation

Eight principal kinds of evidence have been adduced in favor of a spatial representation for information.

Introspective reports. Many subjects in various experiments have reported using spatial imagery to solve transitive inference problems such as linear syllogisms. According to DeSoto et al. (1965), proponents of a spatial imagery theory, "questioning of subjects in Experiment I indicated that most of them solve the syllogisms with the aid of imagery" (p. 516). These investigators also found that different pairs of relational adjectives evoked different kinds of arrays. Better-worse, for example, most often evoked a vertical array constructed in a top-down fashion, whereas darker-lighter most often evoked a horizontal array constructed in a left-right fashion. Huttenlocher and Higgins (1971), also proponents of a spatial imagery theory, found that "Ss report that imagery is intimately tied up with determining the order of items from comparative expressions" (pp. 495-496). Their subjects constructed imaginary spatial arrays by placing the grammatical subject of a sentence in the array, and then moving the grammatical object with respect to the subject. Even Clark (1969b), a linguistic theorist, has found widespread use of imagery by experimental subjects. He reported that "49% of the Ss in Clark [1969a] claimed that they used spatial imagery" (p. 402). Clark thus felt obliged to conclude not that subjects fail to use imagery, but that "it has not been demonstrated that the use of spatial imagery differentially affects the solution of three-term series problems" (p. 402).

Need for spatial array to combine premise information. At some point during the course of problem solution, subjects must comprehend the higher-order relation between the two lower-order relations expressed in the individual premises. Such comprehension is tantamount to making the transitive inference needed to solve the problem. Spatial imagery theorists have specified at a reasonable level of detail how such comprehension can take place. Huttenlocher (1968), for example, argued on the basis of subjects' introspections that subjects create an imaginary array of the two items contained in the first premise, and then use information from the second premise as a basis for joining the third item to the first two items. Linguistic theorists, however, have not specified in reasonable detail how the transitive inference is actually made. As Huttenlocher and Higgins (1971) have pointed out, "Clark [1969a] attempted to [include an account of the comprehension of adjectives and comparatives as well as an account of question answering], while basically ignoring the question of how information from the two premises is combined" (p. 487). Clark (1971) has admitted that the "linguistic theory is not complete. For one thing, it does not fully specify how information from the two premises are [sic] combined" (p. 513). Until a linguistic account of the combination process is proposed, students of transitive inference are perhaps obliged to accept, if only by default, a spatial account of the combination process.

Comparability of data patterns for purported imaginal arrays to those for physical arrays. One of Huttenlocher's main arguments in favor of spatial imagery has been that "the difficulty of solving different forms of [linear] syllogisms parallels the difficulty of arranging real objects according to comparable instructions" (Huttenlocher et al., 1970). A series of experiments has shown that the two types of items do indeed show parallel patterns of data (Huttenlocher, Eisenberg, & Strauss, 1968; Huttenlocher et al., 1970; Huttenlocher & Strauss, 1968). In a typical experiment, Huttenlocher and her colleagues would require subjects to arrange blocks (trucks, human figures, or the like) in a physical array. The terms of the arrangement task would be presented in a variety of forms, each form paralleling a form of linear syllogism. Response times and error rates for the physical arrangement task would then be shown to be highly correlated with response times and error rates for the linear syllogisms task.

Symbolic distance effects. Data reported by Potts (1972, 1974) and by Trabasso and his colleagues (Trabasso & Riley, 1975; Trabasso, Riley, & Wilson, 1975) seem strongly to implicate some kind of spatial process in linear ordering problems. In a typical experiment, subjects are taught a linear ordering of items that takes the form (A, B, C, D, E, F). Subjects are trained only on adjacent pairs of items. If subjects store information in the form of propositions relating the presented pairs of items, then one would expect that upon test,[7] subjects would be able to judge more rapidly the relation between a trained and adjacent pair, say, *C* and *D*, than they would be able to judge the relation between an untrained and nonadjacent pair, say, *B* and *E*. The former pair should be immediately available in working memory, whereas the latter pair should require at least one transitive inference relating *B* to *E* via pairs (*B*, *C*), (*C*, *D*), and (*D*, *E*). In fact, subjects are able to judge the relation between *B* and *E* more rapidly than they are able to judge the relation between *C* and *D*. The further apart the two items are, the easier the judgment turns out to be. This symbolic distance effect is compatible with the kind of "internal psychophysics" proposed by Moyer (1973) and Moyer and Bayer (1976), whereby a spatial analogue representation is constructed for the array, and elements of this analogue representation are compared to one another. In this kind of representation, elements that are at greater distances from one another are more easily distinguishable from one another, and hence easier to interrelate. The symbolic distance effect has generally been considered to be incompatible with linguistic theories of representation.

Serial position effects. In the linear-ordering experiments described above, subjects are trained on all adjacent pairs of items in the linear ordering. Trabasso and his colleagues (Lutkus & Trabasso, 1974; Riley & Trabasso, 1974; Trabasso et al., 1975) have found that errors made during training and retraining exhibit a serial-position effect with respect to position of the pairs in the linear ordering: Maximum errors occur on middle pairs, and fewer errors occur on pairs nearer the ends of the ordering. This serial-position effect is interpreted as prima facie evidence for an underlying spatial array (see Bower, 1971). If subjects learning the pairs of items without constructing an underlying array, that is, learning the list as a set of independent paired associates, then one would expect equal numbers of errors at different places along the linear ordering. At worst, one might observe some negative transfer for all points except the end points of the array, since each other point is learned as both greater (better, faster, etc.) than one point and less (worse, slower, etc.) than another point.

Directional preferences within linear orderings. In many of the adjective pairs used in linear syllogism problems, one adjective of a bipolar pair results in more rapid or more accurate solution than the other. For example, use of the adjectives taller and better result in facilitated performance relative to the adjectives shorter and worse (Handel et al., 1968). These authors have proposed that faster solution for the adjectives taller and better can be accounted for by the facts that (a) taller-shorter is represented along a continuum proceeding from top to bottom and better-worse is represented along a continuum proceeding from right to left, and (b) "people proceed more readily in a downward than in an upward direction, and in a rightward than in a leftward direction" (p. 513). This directional principle (b), together with the principle of end-anchoring described below, seemed to account for "the results of otherwise inexplicable variation in difficulty among linear syllogisms" (DeSoto et al., 1965).

End-anchoring effects. Investigators of transitive inference have repeatedly found end-anchoring effects in their data. End-anchoring effects are observed when it is easier to solve a transitive inference problem presented from the ends of an array inward than it is to solve the problem presented from the middle of the array outward. Consider, for example, the array (A, B, C). A problem with the premises, "A is taller than B; C is shorter than B," should be easier to solve than a problem with the premises "B is shorter than A; B is taller than C," because both premises in the first case are end-anchored, whereas neither premise in the second case is end-anchored. DeSoto et al. (1965) were the first to propose end-anchoring as a principle used in the solution of linear ordering problems, and Huttenlocher (1968) advanced this proposal as well. The two accounts differ slightly, however, in that DeSoto et al. proposed that end-anchoring in either premise could facilitate solution, whereas Huttenlocher suggested that under most circumstances, end-anchoring will facilitate solution only if it occurs in the second premise.

Correlations with spatial visualization tests. Shaver, Pierson, and Lang (1974) have reported correlations across subjects between errors in the solution of linear syllogisms and scores on tests of spatial visualization. These correlations varied in magnitude, but an impressive number of them reached statistical significance. These correlations were interpreted as evidence that spatial imagery is used in the solution of linear syllogisms.

Evaluation of Evidence for Spatial Representation

With eight kinds of evidence converging on the same conclusion, one is tempted to accept the conclusion without further ado. Yet, none of the eight kinds of evidence proves to be conclusive considered either by itself or in conjunction with the remaining kinds of evidence.

Introspective reports. Introspective reports of the use of imagery are common, and are acknowledged even by the most prominent linguistic theorist (Clark, 1969b). A long-standing question in psychology, however, has been whether such reports can be accepted at face value. Many psychological investigators have been reluctant to accept introspective reports as more than suggestive, and the recent data and theoretical framework presented by Nisbett and

Wilson (1977) seem to justify this reluctance. Subjects appear to have little or no conscious access to the processes they use in various kinds of cognitions.

Combination of premise information, symbolic distance effects, serial position effects, and end-anchoring effects. Can a linguistic representation account for any of these effects? The answer appears to be affirmative: A small modification and extension of a linguistic representation suggested by Holyoak (Note 2) will predict all of these effects.

In the proposed representation, information about relations among, say, items A, B, C, D, and E of a linear ordering is expressed in the form of a hierarchy of relations, as shown in Figure 1. In the present representation, un-

---

Insert Figure 1 about here

---

like in Holyoak's, nodes at a given level of the hierarchy can contain overlapping information. At the highest level of the hierarchy, all items are clustered together at a single node, and the items are therefore indistinguishable from each other. At the lowest level of the hierarchy, each item forms its own separate cluster, and each item is therefore distinguishable from every other item. Each intermediate level of the hierarchy permits one additional differentiation of items. For example, the second level, containing two clusters, permits differentiation of A and E; the third level, containing three clusters, permits differentiation of A and D and of B and E; the same principle is applicable at each successive level. When a subject interrogates the hierarchy, he or she is assumed to start at the top node and work downward in the hierarchy toward the bottom node. Subjects are further assumed to use a breadth-first search order between levels of the hierarchy, and a random search order within levels. For example, subjects will always interrogate both ABCD and BCDE before interrogating any of ABC, BCD, or CDE. However, the order in which ABCD and BCDE are searched is assumed to be random. In order to determine whether one element is to the left of (or is higher than, faster than, better than, etc.) another, two nodes at some level of the hierarchy must be found such that each element appears in only one of the nodes, and the first element is to the left of the other.

First, this hierarchy constitutes a unified representation of the five items in the linear ordering. All possible order relations are expressed by the relations among the nodes of the hierarchy. Second, the hierarchy can account for symbolic distance effects. Suppose one needs to determine only whether A is to the left of E. One can make this determination at the second level of the hierarchy, since A but not E appears in the left node and E but not A appears in the right node. To determine whether A is to the left of D (or B is to the left of E), however, one must go to the third level of the hierarchy. In general, the closer together two elements are, the further down in the hierarchy one must go to distinguish them, and hence, the longer the search process is assumed to take. Third, the hierarchy can account for serial position effects in learning. The more relations one needs to learn for a given element, the longer that element should take to learn. Note that A and E each appear at five nodes, B and D each appear at eight nodes, and C appears at nine nodes. Thus, the endpoints are easiest to learn, and items become successively more difficult to learn as they approach the middle of the implicit linear array. Fourth, the hierarchy can account for end-anchoring effects, if one assumes that the relational hierarchy, like virtually any other structure, is more easily constructed from the outside-inward than from the inside-outward.

The hierarchical structure described above is offered as a possible way in which subjects might represent information linguistically, rather than as the actual way subjects do represent information. The sole purpose of describing the structure is to show that a fairly straightforward linguistic representation can ccount for many of the effects that have been believed to necessitate a spatial representation.

Comparability of data patterns for imaginal and physical arrays. Huttenlocher's argument that data patterns for reasoning with purported imaginal arrays are very similar to those for placement with actual physical arrays presents a reasonable case for the analogy between the two kinds of arrays. The argument is strong, however, only if an alternative model not based on such an analogy makes differential predictions. Here is where the problem has lay. DeSoto et al. (1965) and Huttenlocher (1968) originally interpreted their data as presenting a strong case in favor of a spatial model of some sort. At the time, these interpretations seemed quite acceptable. Clark (1969b) later showed, however, that the items used in earlier research did not distinguish between the previous spatial models and his new linguistic model. Negative equative items (with premises of the form "A is not as ___ as B") were needed to distinguish between the two types of models. Clark's (1969a, 1969b) data from such items seemed to Clark to support the predictions of his linguistic model but not of the earlier spatial models. Huttenlocher et al. (1970) then presented new data showing the analogy between results from a physical placement task and a linear syllogisms task. Since her 1968 claim had been only that the two kinds of tasks were analogous, she now felt justified in interpreting Clark's data as supporting rather than refuting a spatial or imagerial model. Clark (1972a) then carried the debate still another step further by presenting data from a physical placement task that did not correspond well to data from the linear syllogisms task. As will be discussed later in the article, however, these data showed only that there exist at least some physical placement tasks that yield patterns of results that do not correspond to patterns of results from the linear syllogisms task. To summarize, there seem to be two problems confronting those who argue for a spatial model on the basis of an analogy between results from physical placement and linear syllogisms tasks: First, the analogy does not always hold up; second, the analogy does not distinguish predictions of a spatial model from those of a linguistic one.

Directional preferences. In general, adjectives that encourage top-down or right-left processing are also those that are linguistically unmarked. As will be shown in the next section, linguistic theory also predicts facilitated processing for these adjectives.

Correlations with spatial tests. Available correlational evidence for spatial representation of information is weak, because although Shaver et al. (1974) have shown convergent validation of the spatial hypothesis, they have not shown discriminant validation with respect to the alternative hypothesis. In other words, errors on the linear syllogism task might well have correlated with scores on tests of linguistic reasoning ability as well as with scores on tests of spatial visualization ability because of the general factor that pervades performance on both types of ability tests. In order to provide a stronger test of the spatial hypothesis, one would have to show high correlations between linear syllogism and spatial test performance coupled with low correlations between linear syllogism performance and linguistic test performance.

Evidence for Linguistic Representation

Three principal kinds of evidence have been adduced in favor of a linguistic representation for information.

Principle of primacy of functional relations. According to this principle (Clark, 1969b), "functional relations, like those of subject, verb, and direct object, are stored, immediately after comprehension, in a more readily available form than other kinds of information, like that of theme" (p. 388). This principle forms the basis for the linguistic representation of information in terms of base strings and underlying deep-structural transformations on these base strings. Clark has not offered any direct experimental evidence to support the principle. He has interpreted Piaget's (1928) account of transitive inference as supporting the principle, though, and has noted that "in solving problems out loud, many children verbalize the underlying base strings of comparative statements directly" (p. 399). Moreover, "the children in Donaldson's (1963) studies often made...errors as a result of their comprehension of propositions as base strings" (p. 399). If this observational evidence is being interpreted correctly, then the evidence supports an underlying, linguistic deep-structural encoding of the premise information.

Principle of lexical marking. According to Clark's (1969b) lexical marking principle, "the senses of certain 'positive' adjectives, like _good_ and _long_, are stored in memory in a less complex form than the senses of their opposites" (p. 389). The "positive" adjectives are referred to as unmarked adjectives, and their opposites (for example, _bad_ and _short_) are referred to as marked adjectives. If a marked and unmarked adjective are placed at opposite ends of a continuum, the unmarked adjective will generally form the name of the scale. Thus, we generally think in terms of scales of goodness and tallness, rather than in terms of scales of badness and shortness. Obviously, we could have a scale of, say, shortness. But to ask how short a person is seems immediately to imply that the person is short, whereas to ask how tall a person is does not imply that the person is tall.

If, as Clark claims, marked adjectives are stored in memory in a more complex form than is needed for unmarked adjectives, one might well expect the encoding of marked adjectives to be more time-consuming than the encoding of unmarked adjectives, and indeed, all studies of transitive inference that have investigated both marked and unmarked adjectives have found longer latencies or more errors associated with items containing marked adjectives than have been found with items containing unmarked adjectives. This evidence therefore seems on its face to support the principle of lexical marking.

Principle of congruence. According to Clark (1969b), "information cannot be retrieved from a sentence unless it is congruent in its functional relations with the information that is being sought" (p. 392). According to this principle of congruence, transitive inference problems in which information from the premises is not congruent with the information sought should take longer to solve than problems in which the information is congruent, since additional time is needed in the former type of problem to establish congruence between the question and the premises. Suppose, for example, the question is "Who is best?" and the answer is A. If A were encoded from a premise such as "A is better than B," then solution should be relatively rapid, since A was encoded in terms of the comparative better and the question asks who is best. Suppose that instead, the relevant premise was "B is worse than A," which, according to Clark, can be expanded to "B is worse than A is bad." This premise does not contain information congruent with the question. The question can be answered only if it is reformulated to read, "Who is least bad?" This reformulation takes additional time, and as an added step, increases the probability of an error on a problem. If, as Clark's (1969a, 1969b) data suggest, people do use the principle of congruence in solving transitive inference problems, then further support is provided for the linguistic model.

Evaluation of Evidence for Linguistic Representation

Principle of the primacy of functional relations. The observational evidence to support the principle of the primacy of functional relations is suggestive at best, and certainly no stronger than subjects' direct introspective reports of spatial imagery. At present, the principle seems to stand more as a presupposition for the remaining two principles than as a principle that is testable in its own right.

Principle of lexical marking. The mere existence of a marking effect does not in itself argue for a linguistic representation for information. As noted earlier, a number of investigators have noticed that the unmarked form of a bipolar adjective pair is generally the form that would be expected to appear at the top of a spatial array. According to Huttenlocher and Higgins (1971), "the unmarked adjective would be toward the top because it designates the presence of a space-occupying property, and the marked adjective toward the bottom because it designates the absence of that property" (p. 497). And since DeSoto et al. (1965) proposed that working from the top down is easier than working from the bottom up in a spatial array, a spatial model could account for the marking effect.

If an adjective pair could be found in which the marked form suggested the top of a spatial array and the unmarked form suggested the bottom of a spatial array, then, according to Clark (1969b), it would be possible to disentangle the spatial and linguistic accounts of the marking effect. Such an adjective pair is found in deep-shallow, where deep, the unmarked adjective in the pair, suggests the lower end of a spatial array. Clark (1969b) has reported that when subjects are presented with linear syllogisms containing the adjective pair, deep-shallow, the standard marking effect is obtained. This result, then, supports the linguistic rather than the spatial account of marking. One must be reluctant to decide between representations, however, on the basis of a single adjective pair, especially one that is so unrepresentative of adjective pairs in general. Another adjective pair, early-late, is reported by Clark (1969b) to show results opposite to those predicted by lexical marking, although the results obtained by Handel et al. (1968) with this adjective pair are consistent with a spatial account of their data.

Principle of congruence. Spatial theorists are skeptical that the available data provide adequate support for the principle of congruence. In a series of recent experiments, Potts and Scholz (1975) obtained a congruence effect under some circumstances but not under others. Clark's (1969b) data provide only weak support for the principle of congruence, and "Clark's 1969a data show that when the answer is in the first premise, and the same adjective appears in both premises, a problem is easier when the adjective in the question does not match that in the premises (24.7% errors) than when it does match (31.4%)" (Huttenlocher & Higgins, 1972, p. 424). On the basis of the data from Clark's two articles, therefore, Huttenlocher and Higgins (1972) retain their "original conclusion that there is no strong evidence for 'congruence'" (p. 424).

Processing of Information in Transitive Inference

The controversy over information processing in transitive inference is much less sharply defined than the controversy over information representation, because neither the spatial nor the linguistic theorists have formulated process models. Instead, the theorists have preferred to formulate their theories in terms of principles. These principles have been used as a basis for differential latency and error predictions in a way that suggests that one or more real-time operations may correspond to each principle. These operations in turn lead to differences among item types in latencies and error rates. The correspondence has remained implicit rather than explicit, however, in the writings of DeSoto, Huttenlocher, and Clark. Johnson-Laird (1972) noticed the correspondence, and constructed process models based upon the spatial theory of DeSoto et al. (1965), the linguistic theory of Clark (1969b), and the operational theory of Hunter (1957). The models seem not to be specified in sufficient detail to permit quantification or simulation. Later in this article, process models are proposed that correspond approximately to the spatial and linguistic models, and these process models are quantified to yield explicit latency predictions. Further discussion of information processing, therefore, is deferred until later in the article.

Effects of Treatment and Subject Variables on Representation and Process

This section considers the effects of mode of problem presentation, relational term, practice, and subject differences on the solution of linear syllogisms.

Mode of Problem Presentation

Transitive inference problems have been presented in a variety of ways. The present discussion will be confined to modes of presentation for linear syllogisms, of which there have been four.

Presentation of whole problem for unlimited time. Hunter (1957) and Clark (1969b) presented subjects with full linear syllogisms, and gave subjects as long as they needed to complete solution.

Presentation of first premise followed separately by second premise and question. Huttenlocher (1968) presented each subject with the first premise of a linear syllogism, and then asked the subject two questions intended to assure that the subject understood the premise. Huttenlocher then presented each subject with the second premise and finally the question. Latencies were recorded beginning after presentation of the second premise, and ending with the subject's response to the question.

Presentation of the first two premises followed separately by the question. Potts and Scholz (1975) presented subjects with the two premises of the linear syllogism, and gave them as long as they needed to process the information contained in the premises. When the subject indicated that he or she was ready, the question was presented separately. Latencies were recorded both for premise processing time and for question processing time.

Presentation of whole problem for ten seconds. Clark (1969a), DeSoto et al. (1965), and Keating and Caramazza (1975) presented subjects with full linear syllogisms for a period of 10 seconds. If a subject was able to solve a problem correctly in this amount of time, his or her response was counted as correct. Otherwise, it was counted as an error.

Effects of mode of presentation. No one has systematically investigated the effects of mode of presentation upon representation and process in linear syllogistic reasoning. Data pertinent to these effects will be presented in the experiments described later in this article. Data very recently collected in my laboratory, however, suggest that the effects of presentation mode are much more complex than anyone has realized, and that differences in presentation mode account for certain discrepancies that appear in the literature on linear syllogistic reasoning. These very recent data will be presented in a separate article (Sternberg, Note 3).

Relational Term

The effects of relational terms (usually adjectives) have been most thoroughly studied by DeSoto et al. (1965) and Handel et al. (1968). Two characteristics of the relational terms have received most attention: differences in directional preference between and within bipolar pairs, and differences in difficulty between and within bipolar pairs.

Directional preferences. The research of DeSoto et al. and of Handel et al. has suggested that subjects tend to order certain relational pairs, such as *better-worse*, *father-son*, and *more-less*, vertically in spatial arrays. *Better*, *father* , and *more* are generally represented at the upper end of each array. Other relational pairs, such as *earlier-later* and *faster-slower*, tended to evoke horizontal spatial arrays, with *earlier* and *slower* at the left end of each array. In still other relational pairs, such as *cause-effect*, *farther-nearer*, and *lighter-*

darker, most subjects are inconsistent in their directional preferences.

Directional difficulties. Handel et al. (1968) tested subjects with problems containing a number of different relational pairs. Although they did not explicitly test differences in item difficulty as a function of spatial direction, it is clear from their data that relational terms for which subjects were inconsistent in their spatial directions were more difficult to process than were relational terms for which subjects were consistent. Within relational pairs, DeSoto et al. (1965) and others have found that items are easier when presented with the adjective or a pair that encourages top-down rather than bottom-up processing, or left-right rather than right-left processing.

Practice Effects

Constant strategy. Most theorists seem to assume that subjects are constant in their strategy: The subjects quickly settle upon a strategy--be it spatial or linguistic--and maintain that strategy throughout their problem solving.

Spatial-to-linguistic strategy change hypothesis. Citing the theory and data of Wood (Note 4), Wason and Johnson-Laird (1972) have proposed that

> the inexperienced subject represents the premises in a unified form (with or without imagery) because this is likely to be the normal practical mode of dealing with the relational information. But by dint of sheer repetition this approach is likely to give way to a purer and more formal strategy geared to the specific constraints of the problem....In short, subjects seem likely to pass from an approach analogous to the IMAGE theory to one analogous to the LINGUISTIC theory. (p. 122)

According to this hypothesis, one would expect subjects to follow a spatial model early during their experience with linear syllogisms, and to switch later to a linguistic model.

Linguistic-to-spatial strategy change hypothesis. Shaver et al. (1974) have proposed a strategy change hypothesis that reverses the sequence described above. They noted that Johnson-Laird (1972)

> hypothesized that imagery is abandoned in favor of a linguistic strategy after practice with three-term series problems. The opposite temporal sequence is indicated by our results, suggesting that in this case at least, imagery provided the "more economical and specialized" strategy. (p. 373)

According to this hypothesis, then, subjects are assumed to follow a spatial strategy early during their experience with linear syllogisms, and to switch later to a linguistic strategy.

Individual Differences

There has been relatively little systematic investigation of individual differences in linear syllogistic reasoning. Burt (1919) was the first to discover that these items distinguished between high and low ability individuals, and used the items on a test of mental ability. Keating and Caramazza (1975) found substantial differences in level of performance on linear syllogisms between high and low ability students in the fifth and seventh grades, and also some evidence of strategy differences. Clark (1969b) reported that 49% of his subjects in the Clark (1969a) experiment reported use of visual imagery; Shaver et al. (1974) reported a figure of 73%. More interestingly, they found an interaction between reported use or nonuse of imagery and reported relative difficulty of visual versus oral problem presentation. Subjects who reported use of imagery were also more likely to report that visual presentation was more difficult than oral presentation, whereas subjects who did not report use of imagery were more likely to report that oral presentation was more difficult than visual presentation. These results are consistent with the hypothesis that reading the problems may have interfered with the use of imagery by the visualizing subjects. Shaver et al. (1974) also reported differential patterns of correlations for men and women between errors in solving linear syllogisms and spatial visualization scores: At least some statistically significant correlations were obtained for men but not for women. Finally, Handel et al. (1968) found that subjects differed widely in their directional preferences for certain relational pairs, although these authors did not undertake a systematic investigation of these individual differences. At the very least, we can say that there is evidence of meaningful individual differences in subjects' solving of linear syllogisms, although the nature of these individual differences needs to be elucidated further. An attempt in this direction is made in the data to be presented later.

THREE INFORMATION-PROCESSING MODELS OF TRANSITIVE INFERENCE

Preview

Three information-processing models of transitive inference will be presented. The three models are applied to (and later tested on) the most widely studied kind of transitive inference problem, the linear syllogism. The three models are a spatial model based upon the DeSoto et al. (1965) and Huttenlocher (Huttenlocher, 1968; Huttenlocher & Higgins, 1971) models, a linguistic model based upon the Clark (1969b) model, and a new linguistic-spatial mixed model. Although the first two information-processing models are based upon previous models, they are not isomorphic to these previous models. In order to make the empirical claims of each model specific, and to facilitate comparisons among models, each of the present models (unlike previous ones) was expressed as (a) an information-processing model in flow-chart form, and (b) a linear mathematical model corresponding to the information-processing model. Quantification of the models permitted more rigorous testing of the empirical claims of the models than has been possible in previous research. Although the first two models are not identical to any previous models, they do seem to capture many of the major intuitions of the previous models upon which they are based. The third model is not based upon any particular previous model, although parts of this new model draw upon previous models. In its mixture of linguistic and spatial processes, it seems most akin to ideas that have been proposed by Trabasso (1975).

The particular realizations of the spatial and linguistic models presented below are obviously not the only possible ones. The linear models probably are, however, the simplest mathematical models that could do justice to the original conceptualizations. I have experimented with more complex, nonlinear realizations of the original conceptualizations, but have been unable to find any implementations that improved nontrivially the fits to data obtained with the linear models. Thus, although other, more complex realizations are possible, it remains to be shown that any is better able to account for data than the simple realizations presented here. Similarly, there are numerous possible mixed spatial-linguistic models of linear syllogistic reasoning. Again, however, it remains to be shown that any alternative mixture model is superior to the one presented here. I have been unable to find a superior mixture model--linear or nonlinear--and as will be shown, the residuals of the predicted values from the observed data are generally small and unsystematic.

The models to be presented all agree that there are certain encoding, negation, marking, and response operations that contribute to the latency with which a subject solves a linear syllogism. All full linear syllogisms contain certain terms and relations to be encoded, and require a response. Only some linear syllogisms contain premises with negations and marked adjectives. Although the models agree on the presence of these operations, they disagree as to which of the operations are spatial and which are linguistic. The models also disagree as to what further operations are required. This divergence is particularly important, because it provides the basis for distinguishing among information-processing models via the linear models. Because the models are partially nonoverlapping in the operations alleged to be used in solving linear syllogisms, the models make different latency predictions across item types.

The three models of linear syllogistic reasoning will be presented with reference to an example of a relatively difficult linear syllogism: *C* is not as tall as *B*; *A* is not as short as *B*; Who is shortest? The correct answer is *C*, and by convention, *A* will always refer to the extreme item at the unmarked end of the continuum, and *C* to the extreme item at the marked end of the continuum.

Process Models

Spatial Model

A flow chart for a spatial model is presented in Figure 2. In this and subsequent flow charts, each box represents a mental operation consuming real time. Latency parameters corresponding to each operation are indicated next to each box, but will not be discussed until later.

Insert Figure 2 about here

Solution starts by the subject's reading the first premise, "*C* is not as tall as *B*." The nature of the next operation depends upon whether or not the adjective in the first premise is marked. Since *tall*, the adjective encountered in the first premise of the example, is unmarked, seriation occurs in the preferred, top-down direction: A spatial array is constructed in which C is placed above B.

Next, the premise is checked for the presence of a negation. If a negation is present, the two terms in the spatial array are flipped in space, so that the top term becomes the bottom term and vice versa. In the present example, the positions of terms C and B are reversed so that B is now placed over C.

The subject then reads the second premise. Since the adjective in the second premise, short, is marked, the two terms are placed into a spatial array in the nonpreferred, bottom-up direction, with A below B. Since there is also a negation in this premise, A and B are flipped around in the array, so that B is below A.

Previous work (for example, DeSoto et al., 1965; Trabasso & Riley, 1975) has suggested that end-anchored premises are easier to process than are premises that are not end-anchored. In other words, subjects prefer to work from the ends of the array inward, rather than from the middle of the array outward. A possible reason for this preference is that end-anchored premises bring one to the pivot, or middle term of the series. If one ends up on the middle term, it is immediately available for use as the pivot of the array. If one does not end up on the middle term, one must search for it, taking additional time. In the present formulation of the model (as in DeSoto et al.'s, 1965), end-anchoring facilitates processing of both premises. Huttenlocher (1968) has proposed that in general, only end-anchoring of the second premise facilitates performance, although the difference between the two formulations proved inconsequential to the rank order of the models to be described.

The subject is now ready to seriate the terms from the two premises, combining information from them. According to DeSoto et al. (1965), premise information is easier to seriate from the top down rather than from the bottom up. Therefore, the problem is easier if the top and middle terms of the array occur in the first premise than if the bottom and middle terms occur in the first premise. In the present example, the bottom and middle terms occur in the first premise (C and B respectively), so that the information from the two premises is seriated bottom-up (or in general, in whatever is the more difficult direction for a given array).

Next, the subject reads the question. If the question adjective is marked, the subject seeks the response at the nonpreferred (usually bottom) end of the array. Otherwise, he or she seeks the response at the preferred (usually top) end of the array. In the present example, the question adjective is marked, so that the more time-consuming process is required. The subject has now found the answer to the problem, and can respond.

Linguistic Model

A flow chart for a linguistic model is presented in Figure 3. The subject begins solution by reading the first premise. If the adjective is marked, then the relation represented by the adjective is more difficult to encode linguistically, and additional time is consumed in the encoding. In the example, the adjective tall is not marked, so more rapid encoding is possible. Regardless of whether or not the adjective is marked, information about the relation is stored as a pair of deep-structural propositions: (C is tall+; B is tall). Next, the subject checks for a negation. If a negation is present, then the roles of the terms in the propositions are reversed: (B is tall+; C is tall).

Insert Figure 3 about here

This reversal operation, like the encoding operation for marked adjectives, is a linguistic one. The entire relation is stored in long-term memory. The relation is also stored in working memory, but in compressed form: (B is tall+). The subject compresses the relation because storing the full relation in working memory would use up more processing space than the subject has available to allocate to the relation, given that he or she must retain sufficient space for processing the second premise.

Next, the subject reads the second premise. Here, the adjective short is marked, so additional time is spent in encoding the relation. The relation is initially stored in working memory as (A is short+; B is short), and then, since there is a negation, the roles of the terms are reversed: (B is short+; A is short). Since the heavy space-using encoding operations have been completed, there is no need to compress the second premise. Moreover, since B appears in both premises, it is easily recognized as the pivot term. In this particular case, the pivot is immediately available.

Under some circumstances, the pivot is not immediately available. Suppose the first premise had been "B is not as short as C." Eventually, the subject would have retained in working memory the compressed proposition, (C is short+). But since the C term does not appear in the second premise, the subject is unable to determine from his or her encodings of the two premises what the pivot term is: Each name appears in working memory just once. The subject is assumed to retrieve from long-term memory the missing link between the premises: (B is short). Once this proposition is in working memory, the subject can recognize B as the pivot term. But this pivot search operation has taken additional time. Note that in this model, the search for the pivot is a search for linguistic information, whereas the search was for spatial information in the spatial model.

Having found the pivot, the subject is ready to read the question. (In this model, there is no seriation operation intervening between pivot search and question reading, because subjects are assumed to store separately the functional relations underlying the two premises.) If the question contains a marked adjective, additional time is spent in encoding it. Finally, the subject is ready to solve the problem. In the example, the subject seeks the individual who is shortest. All propositional information is now made available to working memory for the final search. Solving the problem entails finding the individual who is shortest, that is, who is short+ relative to the pivot, but no such individual is found in the example. The problem is that the form of the question is incongruent with the way in which the answer term has been encoded. Whereas the shortest term, C, was previously encoded as tall (relative to the tall+ B), the question asks for the person who is shortest. The subject therefore must make the question congruent with the problem terms as encoded. He or she does so by looking for the least tall individual--someone who is tall- relative to a tall pivot, or tall relative to a tall+ pivot. The subject can now respond with the correct answer, C.

Linguistic-Spatial Mixed Model

Motivation. Two basic ideas motivate the proposed linguistic-spatial mixed model of transitive inference. The first is that in solving transitive inference problems, subjects seem likely to use both linguistic and spatial operations: First they linguistically decode the verbal information presented in the premises; then they spatially recode the information into a form that permits the transitive inference to be made. This kind of mixture model is consistent with the obvious need for subjects to interpret the verbal input presented to them, and with their frequent reports of spatial imagery in combining information from the two premises. The position adopted here is similar to that adopted by Lawson (1977), who in studying linear ordering problems concluded that

> whatever the nature of the representation, the results of this study indicate that two distinct types of information are available in memory: first, information about the holistic idea conveyed by the entire set of sentences, and second, information in propositional form about what sentences were presented. (p. 9)

Lawson suggests that holistic "knowledge of the ordering is represented in a form that is analogical to a visual depiction of the scene (Huttenlocher, 1968)" (p. 8).

The second basic idea is that a major but previously unappreciated source of difficulty in solving transitive inference problems is the need of the subject at various points in the solution process to locate specific items in the spatial array, in particular, the pivot and the response. In solving a transitive inference problem, the subject's mind's eye traverses the spatial array as necessary. Every time it moves from one location to another, real time is consumed. This notion of visual scanning of a spatial representation is consistent with the sorts of visual scanning processes suggested by Shepard (Cooper & Shepard, 1973; Shepard & Metzler, 1971) and by Kosslyn (1975). The basic notion is that scanning of a visual array is analogous to scanning of a physical array, and in its course consumes measurable time.

Processing strategy. A flow chart for the proposed linguistic-spatial mixed model is presented in Figure 4. The subject begins solution by reading the first premise. In order for the premise to be understood, it must be formulated in terms of the kind of deep-structural propositions proposed by the linguistic model. Encoding a marked adjective into this deep-structural format takes longer than encoding an unmarked one. Also, the presence of a

Insert Figure 4 about here

negation requires a reformulation of the deep-structural proposition. Thus, "C is not as tall as B" is originally formulated as (C is tall+; B is tall), and is then reformulated as (B is tall+; C is tall), as in the linguistic model. Once the deep-structural propositions for the premise are in final linguistic form, the terms of the propositions are seriated spatially. If there is a marked adjective, the subject takes additional time in seriating the relation spatially in the nonpreferred (usually bottom-up) direction. If the adjective is not marked, then the premise is seriated in the preferred (usually top-down) direction. Note that whereas a negation is processed linguistically, a marked adjective is processed first linguistically (in comprehension) and later spatially (in seriation). After seriating the first premise, the subject repeats the steps described above for the second premise.

In order for the subject to combine the terms of the premises into a single spatial array, the subject needs the pivot available. The pivot is either immediately available from the linguistic encoding of the premises, or else it must be found spatially. According to the mixed model, there are two ways in which the pivot can become available immediately: (a) It is the single repeated term from all previous linguistic encodings; or (b) it is the last term to have been linguistically encoded. These rules have different implications for affirmative and negative premises.

In problems with two affirmative premises, the pivot is always immediately available, since each premise has been linguistically encoded just once. One term, the pivot, is distinctive from the others in that more than one relational tag has been associated with it, one from its encoding in the first premise, and one from its encoding in the second premise. The other two terms each have just a single relational tag associated with them. The second principle therefore need not even be applied. Indeed, it is applied only if the first principle fails.

The use of distinctiveness as a cue to the identity of the pivot fails in problems with at least one negative premise. In these problems, each premise containing a negation is encoded in two different ways—in its original encoding and in its reformulated encoding in which the roles of the terms have been reversed. The pivot is therefore no longer the only term with more than one relational tag associated with it, and it thus loses its distinctiveness. The subject must therefore search for the term with the largest number of relational tags, unless he or she can apply the second principle.

When the distinctiveness principle fails, the subject attempts to link the first premise to the last term to have been encoded in working memory. If this term of the second premise happens to be the pivot, the link is successful, and the subject can proceed with problem solution. Pivot search can thus be avoided if the last term to have been encoded is the pivot. But if this term is not the pivot, the link cannot be made, and the subject must search for the pivot--the term with the largest number of relational tags. This search for the pivot takes additional time.

Once the pivot has been located, the subject seriates the terms from the two spatial arrays into a single spatial array. In forming the array, the subject starts with the terms of the first premise, and ends with those of the second premise. The subject's mental location after seriation, therefore, is in that half of the array described by the second premise. The subject next reads the question. If there is a marked adjective in the question, the subject will take longer to encode the adjective, and to seek the response at the nonpreferred (usually bottom) end of the array. The response may or may not be immediately available. If the correct answer is in the half of the array where the subject just completed seriation (his or her active location in the array), then the response will be available immediately. If the question requires an answer from the other half of the array, however, the subject will have to search for the response, mentally traversing the array from one half to the other and thereby consuming additional time.

One final search operation is used optionally under special circumstances. If the subject has constructed a sharp spatial encoding, then he or she is now ready to respond with the correct answer. If the subject's encoding is fuzzy, however, the subject may find that he or she is unable to respond with a reasonable degree of certainty. The subject therefore checks his or her tentative response as determined by the spatial representation with the encoding of that response term in the linguistic representation. If the question and response are congruent, the check is successful, and the subject responds. If the question and response are not congruent, however, the subject reformulates the question to ascertain whether it can be made congruent with the response. Only then does he or she respond.

The notion of optional search for congruence depending upon quality of encoding makes a strong prediction: that the use of this additional operation should be associated with reduced encoding time. Indeed, subjects seek to establish congruence only because they did not take the time to create a sharply defined spatial encoding. Experimental manipulations can be controlled so that the subject is either encouraged to or discouraged from creating a sharp spatial encoding. The experimental manipulations that result in one or the other kind of encoding will be described later, as will the degree to which the data conform to the prediction made above.

Comparison among Models

The models all agree that marked adjectives and negations should increase solution latency. They disagree, however, as to why solution latency is increased. According to the spatial model, solution latency is increased because processing of negations and marked adjectives requires a more complex encoding of information into a visualized spatial array. According to the linguistic model, the additional time results from increased difficulty in a linguistic encoding process. According to the mixed model, negations require a more complex linguistic encoding process, whereas marked adjectives require first more complex linguistic encoding and then more complex spatial encoding.

The models also agree that some form of pivot search is needed under special circumstances. The models disagree, however, as to what these circumstances are. In the spatial model, pivot search is required for premises that are not end-anchored, that is, for premises in which the first term is the middle rather than an end of a spatial array. Absence of end-anchoring necessitates a search through the visualized spatial array. In the linguistic model, pivot search results from compression of the first premise in the deep-structural encoding. If the term that was dropped from working memory in compression happens to be the pivot term, then the subject has to retrieve that term back from long-term memory. In the mixed model, pivot search is required if the reformulated deep-structural version of a negative second premise does not have the pivot in its latter (and hence most recently available) proposition.

The spatial and mixed models agree that the terms of the two premises are combined into a single, unified representation. This combination is accomplished through a seriation operation in which each of the two partial spatial arrays is unified into a single array. The linguistic model disagrees: Functional relations from the two premises are stored separately.

The linguistic and mixed models agree in the need for an operation to establish congruence between question and answer, but in the mixed model, the establishment of congruence is optional. It is used only when the spatial encoding of terms is of insufficient quality to permit the subject to respond to the problem with a reasonable degree of certainty. No operation for the establishment of congruence exists in the spatial model.

In the spatial model, subjects are hypothesized to prefer working in a certain direction (usually top-down) between as well as within premises. Generally, this preference means that extra time will be spent in seriation if the term at the preferred end of the array does not occur in the first premise. No corresponding "additional latency" exists in either the linguistic or mixed model.

In the linguistic model, subjects search the deep-structural propositions for the term that answers the question. In a spatial array, it is obvious which term corresponds to which question adjective. For example, the tallest term might be at the top, the shortest term at the bottom. In linguistic propositions, there is no such obvious correspondence, so that the subject must check both extreme terms relative to the pivot, seeking the correct answer.

In the mixed model, subjects have to search for the response to the problem if their active location in their final spatial array is not in the half of the array containing the response. Subjects mentally traverse the array to the other half, looking for the response. No corresponding operation exists in either the spatial or linguistic model.

Finally, the models agree that the final operation is a response process, whereby the subject selects his or her answer.

Mathematical Models

Mathematical models were formulated from the information-processing models by assuming that (a) each operation represented by a box in the flow charts contributes toward the total real time consumed in the solution of a linear syllogism, and (b) these contributions toward solution time are additive. The formulation of the mathematical models in the context of the present experiments can be described only after the experiments are described, and so details of the quantification will be deferred until a later section of the article.

EXPERIMENTAL TESTS OF THE MODELS

Four experiments were conducted. The experiments were intended to address the three theoretical questions raised near the beginning of the article. Thus, the experiments provide evidence concerning (a) the representations of information during the solution of transitive inference problems, (b) the processes acting upon these representations, and their individual latencies, and (c) the effects of mode of problem presentation, adjective, session, and individual differences upon representations and processes.

In the first experiment, a precueing paradigm was used in order to separate mathematical parameters that otherwise would have been confounded (see Sternberg, 1977b, Chapter 4). In a precued condition, subjects would receive only part of the problem. They would be asked to do as much information processing as possible on this part of the problem before being shown the problem in its entirety. This precueing paradigm presumably forced subjects to read the question last, whereas Johnson-Laird (1972) has suggested that subjects may typically read the question first, prior to reading the premises. In a second experiment, therefore, a similar precueing paradigm was used, except that the question was presented first, followed by the premises. The precueing paradigm in this experiment, however, like that in the first experiment, may have forced subjects to adopt a linear processing strategy that reflected the particular order in which the lines of the problem happened to be presented during precueing. In a third experiment, therefore, precueing of the sort used in the first two experiments was abandoned. Instead, subjects received both two-term and three-term series problems on separate trials; the combination of item types permitted separation of parameters in much the same way that precueing did. In this experiment, and in the two preceding it, each subject received every problem type with each of three different adjective pairs. But this design may have enabled subjects to recognize the applicability of a particular strategy to one adjective pair, and then to carry this strategy over to other adjective pairs, whereas the subject might never have used this strategy on either of the other adjective pairs had each been presented in isolation. For example, a seemingly spatial adjective pair like taller-shorter might prime a spatial strategy, whereas a less spatial adjective pair like better-worse might prime a linguistic strategy. Shaver et al. (1974) have argued that priming does indeed occur. In the fourth experiment, therefore, the procedures of Experiment 3 were repeated, except that each subject received items, all of which contained the same adjective pair.

Method

Subjects

Subjects in Experiment 1 were 16 Stanford undergraduates preselected from the introductory psychology subject pool. All 441 subjects in introductory psychology received two brief ability tests: (a) a 3-minute word-classification test requiring the subjects to select one of five words that didn't belong with the other four; (b) a 3-minute mental rotation test requiring subjects to identify which of a number of geometric forms were rotated versions of a target ("same") and which were both rotated and reflected versions of the target ("different"). The tests were administered to all students simultaneously, and were of the pencil-and-paper variety. Four subjects were then selected for each of four groups: Hi Verbal - Hi Spatial, Hi Verbal - Lo Spatial, Lo Verbal - Hi Spatial, Lo Verbal - Lo Spatial. A high score was defined as one between the 75th and 95th percentiles on a given test; a low score was defined as one between the 5th and 25th percentiles.

Subjects in each of Experiments 2 and 3 were 18 Yale students, and in Experiment 4 were 54 Yale students, from introductory psychology who volunteered to participate in order to receive credit toward a course participation requirement. Subjects were nonoverlapping between experiments, and were not prescreened in any way.

Materials

Stimuli. The basic experimental stimuli were 32 types of linear syllogisms. Items were constructed by varying whether (a) the adjective in the first premise was marked or unmarked, (b) the adjective in the second premise was marked or unmarked, (c) the adjective in the question was marked or unmarked, (d) both premises were affirmative (for example, John is taller than Bill) or negative equative (for example, Bill is not as tall as John), (e) the correct answer was in the first premise or the second premise. In Experiment 1, three adjective pairs were used: taller-shorter, older-younger, faster-slower. In Experiments 2, 3, and 4, the adjective pair better-worse was substituted for older-younger. Terms of the problems were common first names. Half of the names were of men and half of women, although men's and women's names never both occurred in the same problem. Half of the names were one syllable in length; half were two syllables in length, although all names within a given problem had the same number of syllables.

Ability tests. In Experiment 1, subjects received three verbal tests, three spatial tests, and two abstract reasoning tests. The verbal tests were synonyms-antonyms and verbal analogies from the Concept Mastery Test, and the word classification test used for preselection. The spatial tests were the Minnesota Paper Form Board, the French Cube Comparison Test, and the French Card Rotation Test, which was the test used for preselection. The latter two tests are from the French Kit of Reference Tests for Cognitive Factors (French, Ekstrom, & Price, 1963). The abstract reasoning tests were Figural Analogies from the 1934 form of the American Council on Education (ACE) Psychological Examination for College Freshmen, and the French Figure Classification Test. An additional test, the Gordon Test of Visual Imagery, was also used, but since it was uncorrelated with anything else, it will not be considered further.

In Experiment 2, two more verbal tests, a second verbal analogies test and a sentence completion test, were added to the battery of tests described above. The Gordon Test was deleted from this and subsequent experiments. Test 2 (figure classification) from the Cattell Culture-Fair Test of g, Form A, was substituted for the French Figure Classification Test. Otherwise, the same tests were used as in Experiment 1. In Experiment 3, the Concept Mastery Test was deleted, but all other tests were the same as in Experiment 2. Experiment 4 used the sentence completion test, Form S of the Differential Aptitude Test (DAT) Verbal Reasoning (analogies) subtest, French Card Rotation, French Cube Comparison, Form S of the DAT Abstract Reasoning (series) subtest, and the ACE figural analogies.

Apparatus

In Experiment 1, linear syllogisms were presented via an Iconix three-channel tachistoscope. In Experiments 2-4, linear syllogisms were presented via a Gerbrands two-channel tachistoscope. Timing was to the nearest millisecond.

Procedure

Experiment 1. Subjects were first told the nature of the linear syllogism, and then were introduced to the tachistoscope and how to use it. Finally, they were informed of the manner in which the linear syllogisms would be presented. They were told that each trial would be divided into two parts: precueing and solution. In the first part of the trial, subjects might be presented with advance information that would help them solve the problem. Subjects were told to do as much processing as possible on this advance information, taking as long as they needed but no longer to utilize the information fully. They were then to press a foot pedal, which would result in the full linear syllogism appearing on the screen. They were to solve each problem as rapidly as possible without making an error. They were then to indicate by pressing one of three buttons on a button panel which of three responses (left, middle, right) was correct.

The precueing manipulation was similar to that used by Potts and Scholz (1975) in their study of linear syllogisms, and by Sternberg (1977a, 1977b) in the study of analogies. There were two conditions of precueing. In the first, only a lighted blank field appeared. This condition, of course, supplied no advance information. In the second condition, the two premises of the problem appeared, for example, "Sam is taller than Joe. Joe is taller than Bob." The full problem always appeared in the second part of the trial. A typical problem, typed in IBM ORATOR typeface on a 6 x 9-inch index card, appeared in the following form:

Sam is taller than Joe.

Joe is taller than Bob.

Who is tallest?

Joe Bob Sam

Names appeared on the bottom line in random order.

Testing was done over two sessions. The first session began with two trial blocks of 16 items each, half cued and half uncued. Test blocks also consisted of 16 items each, with testing alternating between cued and uncued items. In this experiment, adjective was confounded with presentation order. Subjects received items with the adjective pair taller-shorter followed by faster-slower in Session 1. In Session 2, subjects received blocks with both of these adjective pairs plus the pair older-younger. In this session, there were twice as many items with the pair older-younger as with the other two pairs, so that by the end of the session, subjects had received each of 32 item types with each adjective in each cue condition. Ability testing was distributed over the two sessions, with the ability tests always administered at the end of each session.

Experiment 2. The procedure in Experiment 2 was the same as that in Experiment 1, with the following exceptions.

First, items were presented in a question-first format. All items appeared typed in IBM ORATOR typeface on 4 x 6-inch index cards in the following form:

Who is tallest?

Sam is taller than Joe.

Joe is taller than Bob.

Joe Bob Sam

Second, there were three rather than two precueing conditions. In the uncued (zero-cue) condition, subjects received a blank field in the first part of the trial. In the one-cue condition, subjects received just the question in the first part of the trial, for example, "Who is tallest?" In the two-cue condition, subjects received both the question and the premises in the first part of the trial, for example, "Who is tallest? Sam is taller than Joe. Joe is taller than Bob." The full item was always presented in the second part of the trial.

Third, the adjective pair better-worse was substituted for older-younger, with taller-shorter and faster-slower remaining as in Experiment 1. Furthermore, the presentation of adjectives was completely counterbalanced over three sessions. Again, each subject received each item type with each adjective in each precueing condition.

Fourth, the first part of the trial was terminated by a button rather than a foot pedal.

Experiment 3. In Experiment 3, as in Experiment 1, items were presented with the question following the premises. The choice of three adjectives was the same as in Experiment 2, with presentation of adjectives over three sessions completely counterbalanced.

In this experiment, there was no precueing. Subjects received only full problems. But in addition to receiving linear syllogisms (three-term series problems), subjects also received two-term series problems, which took the following form:

Sam is taller than Joe.
Who is tallest?
Sam Joe

Following Clark (1969b), the ungrammatical superlative rather than the grammatical comparative was used in the question in order to increase uniformity with the linear syllogisms. Order of names on the bottom line was random. Each subject received each three-term series problem type three times with each adjective; each subject received each two-term series problem type four times with each adjective. The identical items were never used more than once, however.

Names were changed on repetitions of problem types. The eight types of two-term series problems used in the experiment varied in whether (a) the premise adjective was marked or unmarked, (b) the question adjective was marked or unmarked, (c) the premise was affirmative or negative equative.

Experiment 4. Procedure in Experiment 4 was the same as in Experiment 3, except that (a) 18 subjects in each of three groups received items with only one of the three adjective pairs taller-shorter, better-worse, faster-slower, (b) each subject received each two-term series problem type three times with each adjective pair, (c) testing was done in two rather than three sessions, with the first session devoted to series problems and the second session devoted to ability testing.

Design

Experiments 1, 2, and 3 used fully within-subjects designs: All subjects received all items. Experiment 4 used a between-subjects design across adjective types: Subjects received problems with only one of the three adjective pairs. In Experiments 1 and 2, precueing conditions were completely crossed with item types, that is, each of the 32 types of linear syllogisms appeared in each cueing condition. In these experiments, the dependent variables were response times for the first and second part of each trial. In Experiments 3 and 4, the single response latency for each item was the dependent variable. Independent variables were adjective markedness for each premise and the conclusion, polarity of premises (affirmative or negative equative), and location of correct response (first premise or second premise).

Quantification of Information-Processing Models

Parameters

Parameters representing the duration of each information-processing component are shown next to each box of the flow charts. The design of the experiments made it possible to estimate some, but not all, of these parameters in an unconfounded fashion. Table 1 shows the parameters that were estimated for each model in each experiment. The contents of the table will be interpreted fully for the mixed model. Interpretation for the other two models follows along the same lines.

Insert Table 1 about here

The design of the experiments made it possible to estimate the durations of negation (NEG), pivot search (PSM), response search (RS), and noncongruence (NCON) in an unconfounded fashion in each experiment. The optional noncongruence operation was relevant to the task only in Experiments 3 and 4, where the absence of precueing was hypothesized to result in hastier and less sharp encodings. The NCON parameter was thus estimated only in the last two experiments. Response component time (RES) could be estimated in isolation only in Experiment 2. This parameter includes time to scan the presented answer options, as well as to indicate a response.

In all four experiments, seriation of the two premises into a single array (SER), premise reading (PR), and encoding and seriation of relations described by unmarked adjectives (NMAR1 and NMAR2) were confounded. The single estimated parameter for these confounded encoding operations has been designated ENC+. (The "+" in this and future parameters represents a mixture of operations.) ENC+ contains a slightly different mixture of operations in Experiments 1 and 2 ($ENC+_1$)

from that in Experiments 3 and 4 ($ENC+_2$). The marking parameter, MARK, was estimated as incremental time for processing marked adjectives over time for processing unmarked ones. Additional time for linguistic encoding of relations expressed by marked adjectives over time for relations expressed by unmarked adjectives is equal to MARK1-NMAR1. Additional time for spatial seriation of relations expressed by marked adjectives is equal to MARK2-NMAR2. In Experiment 2, question reading time (QR) was estimated as the confounded QR+, since it included small amounts of encoding and seriation time for unmarked adjectives (NMAR1 and NMAR2). Response component time (RES) was confounded in Experiment 1 with question reading time (QR) and with some time for encoding and seriating of unmarked adjectives (NMAR1 and NMAR2); in Experiments 3 and 4, it was also confounded with some premise reading time (PR). The confounded parameter for uncued and cued conditions combined is designated RES+. In the uncued conditions alone, it was not possible to separate encoding from response operations, leading to additional confoundings for uncued data analyzed separately. The confounded parameter estimated from uncued data only is designated RES++, and is the sum of (ENC+)+(RES+).

In all, six parameters were estimated for the mixed model in Experiment 1, using the combined uncued and cued data. Seven parameters were estimated for the mixed model in Experiments 2, 3, and 4, again using the combined cued and uncued data. (In Experiments 3 and 4, two-term series problems served the same function as precueing in Experiments 1 and 2--that of separating additional parameters.) Six parameters were estimated for the linguistic and spatial models in Experiments 1, 3, and 4, and seven were estimated in Experiment 2. Further precueing (for example, separation of the first premise of a linear syllogism from the remainder of the problem) could have been used to separate further some of the still confounded parameters, but the additional information to be gained did not seem to justify the increase in the number of parameters that would need to be estimated.

Parameter Estimation

Parameter estimation was done by linear multiple regression, using solution latency for each item type as the dependent variable, and structural aspects of the items as independent variables. Solution time was predicted as the sum of the number of times each hypothetical operation had to be executed, which was given as an independent variable, times the duration of each hypothetical operation, which was estimated as a parameter. Structural aspects and values of the independent variables derived from them are shown for three-term series problems in Table 2, and for two-term series problems in Table 3. At the left side of each table, a shorthand notation is used to describe each item type. The symbol " > " is used to denote an unmarked adjective relating two terms; the symbol " < " is used to denote a marked adjective. A slash drawn through either of these two symbols, " ≯ " or " ≮ ", denotes the expression "not as ___ as."

Insert Tables 2 and 3 about here

The body of each table shows the values of the independent variables used as multipliers to estimate the parameters designated at the top of each column. Values are not shown for the conditions with precueing in Experiments 1 and 2, although these values can be easily inferred (as discussed below). Three parameters, encoding time (ENC+), incremental marking time (MARK), and negation time (NEG), were estimated in the same way for each model (that is, the independent variables were identical), although these parameters have different information-processing implications in each model (see Figures 1-3) and may contain different mixes of hypothesized component processes (see Table 1). Pivot search time (PSM) and response search time (RS) were also estimated for the mixed model, as well as noncongruence time (NCON) in Experiments 3 and 4. Noncongruence time (NCON) and linguistic pivot search time (PSL) were also estimated for the linguistic model. Spatial pivot search time (PSS) and incremental seriation time for the nonpreferred direction (SERN) were also estimated for the spatial model. SERN is the time it takes to seriate in the nonpreferred direction as an increment over time taken to seriate in the preferred direction.

Perusal of Table 2 will reveal that the value of ENC+ remains constant over all 32 types of three-term series problems. It was for this reason that either precueing (Experiments 1 and 2) or two-term series problems (Experiments 3 and 4) were also needed in order to estimate this parameter. These additional problem types also provided other bases for estimating parameters beside ENC+, as can be seen in Table 3 for the case of the two-term series problems.

The use of precueing affected the values of the independent variables in systematic ways. Consider the effects of precueing upon the values of independent variables for Experiment 1. Recall that subjects were presented with the premises in the first part of the trial, so that they needed to process only the question and response alternatives in the second part of the trial. In the cued condition, the independent variable for encoding (as shown in the ENC+ column) drops to 0 for all 32 item types, since all elements entering ENC+--seriation, premise reading, and processing of unmarked premise adjectives--are assumed to have taken place during the first part of the trial. The number of marked adjectives is always 0 or 1, depending upon whether or not the question adjective is marked. The number of negations is always 0, since negations occur only in the premises. Pivot search becomes irrelevant (and hence of value 0) in each model, because it is confined to processing of the premises. Incremental seriation in the nonpreferred direction also becomes irrelevant in the spatial model. The response search and noncongruence processes remain, however, with the same values as for the uncued items, since these processes cannot take place until the question has appeared, and the question does not appear until the second part of the trial.

In all, there were 32 uncued item types in each experiment. In Experiment 1, there were an additional 32 cued item types, so that a total of 64 data points had to be predicted. In Experiment 2, there were an additional 64 cued item types, for a total of 96 data points. In Experiments 3 and 4, there were an additional 8 two-term series item types, for a total of 40 data points.

Results

Preview

The results of the experiments will be presented in five major parts. First, basic statistics for the linear syllogisms data will be presented. Second, qualitative aspects of the fits of the mathematical models to the latency data will be described. Third, quantitative aspects of the fits of the models to the data will be described. Fourth, the latencies of individual component processes in linear syllogistic reasoning will be discussed. Fifth, individual differences in transitive inference will be examined, and will be shown to help distinguish spatial from linguistic processes.

Basic Statistics

Table 4 presents basic statistics--means and their standard errors over items--for the data sets that were used in mathematical modeling. There was no significant difference across experiments in overall mean solution latencies for uncued items, $F(3,124) = 1.61$, $p > .05$. Because selection of adjectives, numbers

Insert Table 4 about here

of sessions, and precueing manipulations differed across experiments, no significance tests were conducted on these data sets. Visual inspection reveals no consistent effect of adjective, however. In Experiments 1 and 3, there was a monotone decline in solution time across sessions, although this pattern did not appear in Experiment 2.

The error rate in each experiment was 1%.

Qualitative Fits of the Models to the Data

A five-way analysis of variance was conducted on observed solution latencies for uncued items and on predicted solution latencies for uncued items for each of the three mathematical models. The five factors in the analysis were the same ones that generated the $2^5 = 32$ uncued item types used in each experiment: (a) markedness versus unmarkedness of the first premise adjective, (b) markedness versus unmarkedness of the second premise adjective, (c) markedness versus unmarkedness of the question adjective, (d) affirmative versus negative equative premises, (e) presence of the correct answer in the first versus the second premise. Each cell of the 2^5 design contained four observations, namely, the means over subjects of the solution latencies for a given item type for a given experiment. Data for individual subjects will be discussed at length later in the article. In interpreting the results of the analysis of variance, $\alpha = .05$ was used as the minimum level for statistical significance.

Main effects of marking. The observed data showed statistically significant effects for marking of the first premise adjective, $F(1,96) = 48.53$, $p < .001$, marking of the second premise adjective, $F(1,96) = 25.42$, $p < .001$, and marking of the question adjective, $F(1,96) = 20.12$, $p < .001$. All three models predicted these statistically significant marking effects. Whereas the three models predicted the effect of marking to be the same for all three adjective positions, however, the data showed differential effects for the three positions. The predicted effect of marking was 439 msec for each model for each adjective. The observed effects were 556 msec for the first premise adjective, 402 msec for the second premise adjective, and 358 msec for the question adjective. The models, therefore, were all satisfactory in accounting for the existence of a marking effect, but were all unsatisfactory in failing to account for differences in the magnitude of the effect as a function of which adjective was marked.

Main effect of negation. The observed effect of negation, 1086 msec for the two premises combined, was statistically significant, $F(1,96) = 185.12$, $p < .001$. The predicted effect for each of the three models was also statistically significant, and was equal in magnitude to the observed effect, 1086 msec.

Main effect of answer in first premise. Items with the correct answer in the first premise were significantly harder than items with the correct answer in the second premise, $F(1,96) = 55.93$, $p < .001$. The mixed model correctly predicted that items with the correct answer in the first premise would take 598 msec longer than items with the correct answer in the second premise. This statistically significant added latency reflects the need of the subject to search for the response in items where the response is not immediately available. The linguistic and spatial models, however, lacking a response search operation, predict no difference in latency as a function of which premise contains the correct answer.

Interactions. A detailed accounting of interaction effects would consume more space than it warrants. The observed data showed seven statistically significant interactions, of which two were two-way, three were three-way, and two were four-way. The mixed model correctly predicted three of these, the linguistic model, two, and the spatial model, one. Moreover, each model predicted two spurious interactions (although not the same two in each case). It is reasonable to conclude, therefore, that none of the simple linear models does full justice to the complexity of information processing shown by subjects in the solution of linear syllogisms. At best, the simple linear models are approximations to the complex and possibly nonlinear processing strategies subjects use. It remains to be shown, however, that even a somewhat more complex model could substantially improve qualitative fit to the data, especially in view of the complexity of the observed interactions. In correctly accounting for all of the main effects and some of the interactions, the mixed model seems at least to be a good approximation to the true model.

Predicted and observed values for individual data points. The observed data and the predictions of the mixed model for each of the four experiments are shown in Table A of the appendix. An examination of this table reveals that the predicted times show very good agreement with the observed ones. On the other hand, there are some points that are either underpredicted or overpredicted in all four experiments, showing, as did the above analyses, that there is room for improvement in the mixed model.

Quantitative Fits of the Models to the Data

With data for the 32 uncued item types averaged across subjects and experiments, the mixed model accounted for 90.2% of the variance in the data with a root-mean-square deviation (RMSD) of 28 csec; the linguistic model accounted for 67.5% of the variance in the data with an RMSD of 52 csec; the spatial model accounted for 62.9% of the variance in the data with an RMSD of 55 csec. Table 5 presents squared correlations between predicted and observed data points for a number of experimental data sets based upon group means for each experiment. All latencies, including those for error trials, were used in modeling.

Insert Table 5 about here

Overall uncued solution time. The data of primary interest are those from the first data set, representing latencies for all 32 uncued item types averaged over adjectives and sessions. These data will be considered in greater detail than the others.

In each experiment, the mixed model is clearly superior to either the linguistic or spatial model: The differences in R^2 between the mixed model and the second best model, the linguistic one, are .213, .148, .155, and .240 in Experiments 1, 2, 3, and 4 respectively. Thus, regardless of whether the question came before or after the premises, of whether or not precueing was part of the experimental design, and of whether different adjectives were presented within or between subjects, the mixed model best accounted for the data.

In Experiments 1 and 2, the same number of parameters was estimated for each model, so that there is no question regarding comparability of model fits for the mixed, linguistic, and spatial models. In Experiments 3 and 4, the mixed model had one additional parameter, the optional noncongruence parameter. According to the mixed model, this additional parameter is necessary, and any model with fewer parameters than the mixed model is inadequate by virtue of being incomplete. If the additional parameter is deleted from the mixed model, however, the mixed model is still superior to the alternative models. With noncongruence deleted, values of R^2 were .765 in Experiment 3 and .832 in Experiment 4. The differences in R^2 between the mixed model and the linguistic and spatial models respectively are .076 and .188 in Experiment 3, and .121 and .189 in Experiment 4. Thus, even without the optional noncongruence parameter of Experiments 3 and 4, the mixed model retains its superiority over the other models.

The levels of fit for the three models must be assessed in the context of the reliabilities of the data. Reliabilities of solution times for the uncued items were computed in two ways, within experiment and between experiment.

The within-experiment (internal consistency) reliabilities were computed by correlating mean latencies on each of the 32 item types for odd versus even numbered subjects. (Numbers were assigned to subjects in an arbitrary fashion.) These correlations were then adjusted by the Spearman-Brown formula. This formula takes into account the fact that only half of the observations were used in each of the two sets of observations that were correlated with each other. The within-experiment reliabilities indicate the proportion of true or systematic variance in each set of data, and thus set an approximate upper limit on the level of fit (R^2) that any one model can be expected to show. Within-experiment reliabilities were .86, .82, .92, and .99 in Experiments 1, 2, 3, and 4 respectively. Considered in conjunction with the fits of the mixed model, the reliabilities show that in Experiment 1, of .19 unexplained variance, .05 was systematic and .14 was unsystematic; in Experiment 2, of .26 unexplained variance, .08 was systematic and .18 was unsystematic; in Experiment 3, of .16 unexplained variance, .08 was systematic and .08 was unsystematic; in Experiment 4, of .12 unexplained variance, .11 was systematic and .01 was unsystematic.

The between-experiment reliabilities were computed by correlating mean latencies on each of the 32 item types across each pair of experiments. These reliabilities indicate the proportion of shared variance across experiments, and thus set an approximate upper limit on the generalizability of any one model to the four sets of data. Between-experiment reliabilities were .84 between Experiments 1 and 2, .78 between Experiments 1 and 3, .80 between Experiments 1 and 4, .83 between Experiments 2 and 3, .87 between Experiments 2 and 4, and .92 between Experiments 3 and 4. These data suggest that no single model could be expected in all three experiments to account for proportions of variance exceeding the low to mid .80's. Thus, the mixed model could not have done much better across experiments than it did.

None of the models accounted for all of the systematic variance in the data. It is of interest to determine whether the unaccounted for systematic variance is statistically significant relative to the total unaccounted for variance. This determination was made by testing the statistical significance of correlations between pairs of residuals of predicted from observed values. Significant correlations indicate unaccounted for variance that is statistically reliable. Significance was tested both within and between experiments. These correlations are presented in Table 6.

Insert Table 6 about here

Within-experiment comparisons were computed by splitting subjects into odd- and even-numbered groups (with numbers arbitrarily assigned), modeling solution times separately for each group, calculating residuals of predicted from observed values for each group, and then correlating the residuals. Resulting correlations were adjusted by the Spearman-Brown formula, since only half the observations were used in the calculation of each set of residuals. One-tailed significance tests were then applied to the correlations, as shown in Table 6. The mixed model could not be rejected in Experiment 1, although it could be rejected at the .05 level in Experiments 2 and 3, and at the .001 level in Experiment 4. The linguistic model could be rejected at the .05 level in Experiment 1, the .01 level in Experiment 2, and the .001 level in Experiments 3 and 4. The spatial model could be rejected at the .05 level in Experiment 1, and at the .001 level in Experiments 2, 3, and 4. The mixed model thus provides the best mathematical account of the data, although in three of the four experimental comparisons, there is statistically significant unexplained variance.

Between-experiment comparisons were computed by modeling solution times separately for each experiment, calculating residuals of predicted from observed values in each experiment, and then correlating the residuals across experiments. The mixed model could be rejected at the .05 level in comparisons between data for Experiments 1 and 3, and at the .001 level in comparisons between Experiments 1 and 4 and between Experiments 3 and 4; it could not be rejected in any other experimental comparisons. The linguistic and spatial models could be rejected at the .001 level in every experimental comparison. The results indicate the superior generalizability across experiments of the mixed model over the linguistic and spatial models.

Solution times for individual adjectives. Returning to Table 5, we see model fits presented individually for each adjective pair. Since these fits are based upon only one-third as much data as the above fits, the values of R^2 are substantially lower. The results are clearcut, however: The mixed model gives a superior account of the data for every adjective in every experiment. The linguistic and spatial models alternate between second and third place with respect to fit.

Although the mixed model is always superior to the other models, it seems to account better for performance with the adjective pair fast-slow than it does with any other adjective pair. None of the other models consistently show this preference for fast-slow, nor for any other adjective pair; nor does the mixed model show any other clearcut trends for other adjective pairs. It is not clear why the mixed model performs better for the fast-slow pair than for other pairs.

Solution times for individual sessions. The table also presents model fits for individual sessions. Once again, the results are clearcut: The mixed model provides a superior account of the data for every session in every experiment.

These data are of particular interest because they are inconsistent with both strategy-change hypotheses: The mixed model is best regardless of session.[1] Moreover, when one considers only the patterns of fit for the linguistic and spatial models, there is no apparent interaction between level of practice and choice of processing strategy. Thus, a direct test of the strategy-change hypotheses failed to provide confirming evidence for either one of them.

Overall uncued and cued solution times. The model fits described in the preceding sections have been based upon uncued solution times only. When solution times from the precued conditions (or in Experiments 3 and 4, two-term series problems) are combined with solution times from the uncued condition in each experiment, model fits increase dramatically, as shown in Table 5. The increase appears to be due to the large increase in solution-time variance introduced by the separation of the encoding components (ENC+) from other components. This separation is possible only because of the use of precueing or two-term series problems. Because the levels of R^2 are so high, the values of R^2 for the various models are closer together than in previous comparisons. Nevertheless, the mixed model once again provides the best fit to the data from each experiment.

Precueing. It is possible to model precueing times as well as solution times. Precueing times, it will be recalled, are those from the first part of the trial in Experiments 1 and 2. In Experiment 1, the modeled data were based upon times to process just the two premises of each problem. In Experiment 2, the modeled data were based upon times to process the question and premises (but not the answer alternatives). Table 5 shows that in both experiments, the mixed model gave an account of the data that was superior to that given by either the linguistic or spatial model. The two-term series problems of Experiments 3 and 4 did not provide an adequate item set for distinguishing among models. With negation time (NEG), marking time (MARK), noncongruence time (NCON), and response time (RES+) as parameters, it was possible to account for .848 of the variance in the two-term series problem latencies of Experiment 3, and .887 of the variance in the two-term series problem latencies of Experiment 4.

Solution times for individual subjects. Table 7 presents data concerning the performances of the models in predicting individual solution-time data both for uncued items only and for uncued and cued items combined. Models are evaluated with respect to mean R^2 for individual subjects and the number of cases in which each model best fit the data of individual subjects.

Insert Table 7 about here

As in previous analyses, the mixed model gave the best account of the data in each experiment both for uncued items only and for uncued and cued items combined. The mixed model did not give the best account of the data in every individual case, however. With the uncued data, it was best in 73% of the cases; with the uncued and cued data combined, it was also best in 73% of the cases.

In a number of individual cases, the fit of the best model to the data was only trivially better than the fit of the next best model. It is therefore of interest to know in what proportion of the cases one model was clearly superior to any other model. Suppose we decide (arbitrarily) that a practically significant difference in individual model fit is represented by a difference in R^2 of .05 or greater. In what proportion of the cases did one model perform significantly better than any of the others? The data in Table 8 address this question.

Insert Table 8 about here

The number at the top of each cell shows the proportion of cases in which one model performed significantly better than another. Thus, using the .05 cutoff, we find that the linguistic model was significantly better than the mixed model for 8% of the cases, the spatial model for 6% of the cases. On the other hand, the mixed model was significantly better than the linguistic model for 65% of the cases, and significantly better than the spatial model for 72% of the cases. The number in parentheses at the bottom of each cell shows the comparable proportion for a practically significant difference in R^2 of .10. Whichever cutoff is used, the proportion of cases for which the mixed model is inappropriate is quite small, whereas the proportion of cases for which the linguistic or spatial model is inappropriate is quite large. The mixed model is thus not only preferred for the group data, but for the large majority of individual cases as well. Note, though, that individual differences do exist: At least some of the 106 subjects in the four experiments used a strategy that was better approximated by the linguistic or spatial models than by the mixed model.

Latencies of Component Processes

Parameter estimates for mixed model. Parameters were estimated as the unstandardized regression coefficients weighting each of the independent variables shown in Tables 2 and 3. Each parameter is hypothesized to correspond to the duration of one or more component processes, as shown in Table 1. Names of parameters are the same as in Table 1.

Table 9 shows values of the parameters and their standard errors as estimated from various sets of data from each experiment, including uncued solution times, combined uncued and cued solution times, and cue times. Estimates from the first two sets of data are obviously nonindependent.

Insert Table 9 about here

The ENC+ parameter includes a combination of times for between-premise seriation, incremental seriation of marked adjectives in the nonpreferred direction, premise reading, and encoding of unmarked adjectives. The first two processes are hypothesized to be spatial, and to account for most of the estimated time. The second two processes are hypothesized to be linguistic. $ENC+_1$ differed significantly from 0 in both experiments in which it was estimated (1 and 2), and was estimated at about 4650 msec. $ENC+_2$, comprising fewer operations, was estimated at about 3050 msec. It seems unlikely that the small difference in the composition of $ENC+_1$ and $ENC+_2$ (see Table 1) could account for the large difference in estimated values. Rather, it seems most likely that encoding operations were performed more rapidly in Experiments 3 and 4, where $ENC+_2$ was estimated, than in Experiments 1 and 2, where $ENC+_1$ was estimated. This difference is exactly as predicted by the mixed model, according to which encoding should be more rapid and less careful in experimental paradigms leading to the use of the optional noncongruence operation. In Experiments 1 and 2, the use of precueing presumably encouraged subjects to encode the premises fully before indicating readiness to see the question and solve the problem. In Experiments 3 and 4, there was no precueing in which subjects could take as long as they needed to get a sharp spatial encoding. Hence, subjects are likely to have encoded the items more quickly and less sharply, at the expense of needing the extra check for congruence at the end.

It was possible to estimate unconfounded durations of negation, marking, pivot search, and response search times in all four experiments. Estimates of negation time center at about 350 msec, of marking time at about 400 msec, of pivot search time at about 1100 msec, and of response search time at about 500 msec. Question reading time (plus confounded operations) could be estimated only in the second experiment, and appears to be about 400 msec.

Response time is about 800 msec.

For the most part, the group parameter estimates are reasonable and in close agreement across data sets. The two exceptions to this agreement are that negation time is inexplicably low in Experiment 3, and response search time is inexplicably low when estimated for cued and uncued data in Experiment 1.

An examination of parameter estimates for individual subjects, assuming use of the mixed model, reveals that the individual data were considerably less reliable than the group data. In Experiments 1, 2, 3, and 4 respectively, the proportions of statistically significant parameter estimates ($\underline{p} < .05$) were 1.00, 1.00, 1.00, 1.00 for ENC+, .56, .50, .44, .37 for NEG, .50, .78, .61, .41 for MARK, .81, .94, .78, .43 for PS, .31, .83, .33, .39 for RS, .56 (Experiment 2 only) for QR+, and .67 and .24 (Experiments 3 and 4 only) for NCON.

Parameter estimates for linguistic and spatial models. Group parameter estimates for the linguistic and spatial models were also computed, and are useful as a diagnostic for assessing where these models failed to predict the data adequately. These parameter estimates are shown in Table 10. The values are for uncued items only, and are presented separately for each of the four experiments.

Insert Table 10 about here

In the linguistic model, values of the negation and marking parameters differed significantly from zero in all four experiments. The value of the noncongruence parameter was significantly different from zero in Experiments 2, 3, and 4. The value of the linguistic pivot search parameter was significantly different from zero only in Experiment 3. If a new, improved linguistic model is to be formulated, it will have to reconceptualize the role of linguistic pivot search and possibly of noncongruence. The linguistic pivot search parameter obviously fails to carry its weight. The linguistic noncongruence parameter is a strong contributor to the model only in Experiments 3 and 4.

In the spatial model, values of the negation and marking parameters differed significantly from zero in all four experiments. The spatial pivot search parameter was significantly different from zero in Experiments 3 and 4, and the parameter for seriation in the nonpreferred direction was not significantly different from zero in any experiment. If a new, improved spatial model is to be formulated, it will probably have to eliminate the parameter for seriation in the nonpreferred direction. The role of spatial pivot search may also have to be reassessed.

Partitioning of total solution time. By multiplying the estimated latency of each operation by the average number of times it is executed, one can estimate the average amount of time spent on each operation during solution of a typical linear syllogism. Figure 5 shows a partitioning of total solution time for a typical negative equative item in each of the four experiments. A typical affirmative item would differ only in the deletion of the latencies for the negation (NEG) and pivot search (PSM) operations. The partitioning assumes the use of the mixed model.

Insert Figure 5 about here

In all four experiments, encoding operations take by far the largest amount of time, whereas response search and noncongruence (where applicable) take the smallest amounts of time. Encoding operations are about 1 1/2 sec faster in the two experiments in which precueing was not used than in the two experiments in which precueing was used. This difference, again, is consistent with the prediction of the mixed model that spatial encoding should be hastier in the experiments without precueing, leading to a less adequate spatial representation of the relations among terms and the subsequent need for a check of the prior linguistic encoding.

Individual Differences in Transitive Inference

Weights of component processes in accounting for individual differences. The parameter estimates presented in the preceding section provided an indication of how important each operation is in accounting for between-items variance. It is of further interest to know how important each operation is in accounting for between-subjects variance. In other words, one seeks to determine the relative contribution of each operation in generating individual differences in overall solution times. Table 11 addresses this question. It shows for each experiment the standardized regression coefficients obtained when subjects' mean solution times for the 32 uncued item types are predicted across subjects by multiple regression from their individual parameter estimates. (Note that all previous modeling was across item types, not subjects.) Since the parameters were estimated from the data on which the means are based (plus precued data as well), values of R^2 were close to 1 and of no interest.

Insert Table 11 about here

The standardized weights show that the encoding operations (ENC+) contribute by far the most to predicting individual differences in overall solution latencies. Thus, in order to understand why individuals differ in the latencies with which they solve linear syllogisms, one's first order of business is to understand more about the nature of the encoding operations.

Intercorrelations between parameters. Intercorrelations between parameters are shown in Table 12. In order to increase the power of the statistical tests, data were combined across all four experiments. There were a total of 106 subjects in the four experiments combined.

Insert Table 12 about here

It was mentioned earlier that many parameter estimates were not statistically reliable for individual subjects. Because of the unreliability of some parameter estimates, correlations were computed in two different ways. One set of correlations (presented in roman type) is based upon the parameter estimates of all individuals for whom the parameter could be estimated. (Recall that not all parameters could be estimated in all experiments, as shown in Table 1.) A second set of correlations (presented in italic type) is based upon only those parameter estimates that were statistically significant at the .025 level. This relatively stringent level of significance was used because of the large number of parameter estimates involved. Obviously, neither set of correlations is ideal. The first set is attenuated by the inclusion of unreliable estimates. The second set may be biased by the inclusion of only subsamples of the data. The direction of the potential bias is unclear: On the one hand, the parameter estimates retained are for subjects who were most clearly using the mixed model; on the other hand, the range of the parameter estimates is restricted because statistically significant parameter estimates tend to be higher ones. However, it will be shown in this and subsequent analyses that although the magnitudes of the correlations differed under the two procedures, the conclusions to be drawn were practically the same.

In computing correlations with ENC+ and RES+, a dummy variable was held constant that distinguished participation in Experiments 1 and 2 from participation in Experiments 3 and 4. First-order partial correlations were used because these parameters were estimated from different mixes of components in the two sets of experiments, and because according to the mixed model, ENC+ should have a lower latency in the latter two experiments, where spatial encoding is assumed to be less careful. Higher-order partials controlling for membership in each experiment were also tried, but had almost no further effect on the magnitudes of the correlations.

The numbers of subjects for whom parameter estimates were significant at the .025 level were 106 for ENC+, 25 for NEG, 27 for MARK, 41 for PS, 28 for RS, 9 for QR+ (Experiment 2 only), 12 for NCON (Experiments 3 and 4 only). RES+ was estimated as a regression constant, and so no significance test was available. All 106 values were used.

Of the 25 possible sets of intercorrelations represented in the table (which excludes unities in the diagonal), 12 are statistically significant under both correlational procedures, 8 are nonsignificant under both procedures, and 5 show discrepancies. There is thus good agreement between procedures. Of most importance is the large number of statistically significant correlations. These relationships show that many pairs of the various latencies are nonindependent. Such a pattern is what would be expected if some of the processes are essentially linguistic in nature and others are essentially spatial. The one component that is correlated with every other component, ENC+, is a parameter hypothesized to contain a mix of both linguistic processes (premise reading and encoding of relations expressed by unmarked adjectives) and spatial processes (between-premise seriation and within-premise seriation of relations expressed by unmarked adjectives), as shown in Table 1. Marking is also correlated with most of the other parameters, and it too is hypothesized to contain a mix of linguistic and spatial processes.

Correlations between parameters and composite ability scores. Composite ability scores were computed by standardizing scores on each ability test (within experiment), summing these standard scores for each type of test, and then restandardizing the sum. Within-experiment standardization was required because different ability tests were used in different experiments. Although the particular tests varied, the measured abilities were the same: verbal, spatial visualization, abstract reasoning. Verbal items included tasks such as synonyms-antonyms, verbal analogies, verbal classifications (requiring subjects to recognize which one of five words didn't belong with the other four), and sentence completions (requiring subjects to indicate which of five words best fit in a blank embedded in the context of a sentence). Spatial visualization items required mental rotation or rearrangement of geometric forms in two or three dimensions. Abstract reasoning items included geometric analogies, geometric classifications, and geometric series. The particular tests used are named in the Materials section of the Method.

The verbal composite was only weakly correlated with the spatial composite, $r = .20$, $p < .05$, and with the abstract reasoning composite, $r = .21$, $p < .05$. The spatial and abstract composites were highly correlated, however, $r = .65$, $p < .001$, suggesting that the two types of tests measured similar abilities. This is a standard pattern of correlations in the psychometric literature (see Cattell, 1971), and in some tests, such as the Cognitive Abilities Test, spatial and abstract reasoning tests are combined in the computation of a single, nonverbal score.

Correlations between parameter estimates and composite ability scores are

Insert Table 13 about here

shown in Table 13. The correlation between overall solution latency in the uncued condition of each experiment and composite ability score is also shown. Correlations were computed in the same ways described in the preceding section. The two methods of computation showed a high degree of consistency. Of 24 possible correlations, 17 were statistically significant under both methods, 6 were nonsignificant under both methods, and only 1 was significant under one method but not the other.

Overall solution latency was highly correlated with all three types of tests. This pattern of correlations is consistent with the mixed model, since solution of linear syllogisms is hypothesized to require both verbal and spatial-abstract processes. The pattern is not consistent with models that postulate that the solution process is either strictly linguistic or strictly spatial-abstract.

The encoding parameter (ENC+) was also significantly correlated with all three ability composites. This result is consistent with the mixed model, according to which the ENC+ parameter includes both linguistic and spatial-abstract processes. A strictly spatial or linguistic model would have trouble accounting for this pattern. Although ENC+ contains a mixture of operations, the predominant operation, according to the mixed model, is spatial seriation between premises. This is the crux of the three-term series problem, and the major source of difficulty. Hence, the model predicts that the spatial-abstract correlation will predominate, and this is in fact the case. The correlations of ENC+ with both spatial and abstract scores are greater in magnitude than -.5, whereas the correlation with the verbal score (presumably due primarily to premise reading) is only -.25. These data suggest that the premise terms are indeed encoded into some kind of spatial array.

The negation parameter (NEG) shows significant correlations with the spatial and abstract composites but not the verbal composite. This pattern of correlations is inconsistent with the prediction of the mixed model, according to which negation is a linguistic operation. The obtained pattern of correlations suggests that as hypothesized by the spatial model, negation is accomplished spatially by reversal of the positions of the two terms in a within-premise spatial array. The mixed model may have to be revised to reconceptualize negation as a spatial-abstract process. Latency predictions would remain the same.

The marking parameter (MARK) shows some relationship to verbal, spatial, and abstract composites, as predicted by the mixed model but neither the spatial nor linguistic models. The relationship to spatial-abstract ability appears to be substantially stronger than that to verbal ability, suggesting that the primary source of individual differences is in spatial seriation of terms (MARK2) rather than in linguistic encoding of the marked relation (MARK1).

Pivot search (PSM) shows significant correlations with the spatial and abstract composite but not with the verbal composite. This pattern of correlations is consistent with the mixed model, which postulated pivot search to be a spatial-abstract operation. Neither the linguistic nor the spatial model contains the pivot search operation as conceptualized by the mixed model, so no relevant predictions can be made for these models.

Response search (RS) is significantly correlated with all three types of tests. According to the mixed model, however, response search was supposed to be exclusively a spatial process. It now appears that in searching for a response, subjects may differ in the rates at which they read off names from an array as well as in the rates at which they can traverse distances in the array. The two types of individual differences would account for the dual linguistic and spatial correlations (Clark, Note 5).

Noncongruence (NCON) is significantly correlated with the verbal but neither the spatial nor the abstract composites. This correlational pattern is consistent with the mixed (or linguistic) model, which stipulates that noncongruence is an optional linguistic operation.

Finally, response (RES+) is significantly correlated with the verbal composite but not with the spatial and abstract ones. Examination of Table 1 reveals that the response parameter as estimated for the mixed model contains up to three linguistic processes--question reading (QR+), premise reading (PR), and encoding of relations expressed by unmarked adjectives (NMAR1). The parameter contains just one spatial process--seriation within premise of relations expressed by unmarked adjectives (NMAR2). The pure response component (RES) itself is not identified in advance as either linguistic or spatial. The obtained results, therefore, are consistent with the larger number of linguistic operations hypothesized by the mixed model to constitute the response component.

In general, the results of this individual-difference analysis are supportive of the mixed model, according to which particular operations should show patterns of individual differences along either verbal, spatial, or both lines. Two results suggest the need for possible changes in the mixed model. The first is the significant correlation of the negation parameter (NEG) with the spatial and abstract composites but not the linguistic composite. The second result is the small but significant correlation of response search (RS) with verbal as well as spatial and abstract abilities. In the case of negation, the nature of the suggested reformulation is evident, since a spatial account of negation has been suggested whereby terms are flipped in a spatial array. In the case of response search, subjects may differ in the rate at which they read off names in an array.

Correlations of composite solution latencies for individual adjectives and sessions with ability test composites. Table 14 shows correlations between ability test composites and uncued solution latencies for combined data, individual adjectives, and individual sessions. The correlations were computed separately for each experiment, since the choice of adjectives and numbers of sessions differed from one experiment to another. Although correlations of ability scores with parameters for each adjective and each session would also be of interest, the individual subjects' data were not reliable enough to permit exploration of these relationships.

Insert Table 14 about here

The correlations with individual adjectives are of interest in determining whether more clearly spatial adjective pairs, such as taller-shorter, better tap individual differences in spatial-abstract ability than do less obviously spatial adjective pairs, such as better-worse, which would seem more likely to lend themselves to a linguistic strategy. Indeed, Clark's (1969a, 1969b) major support for the linguistic model of linear syllogistic reasoning is based upon data collected for the single adjective pair, better-worse. The possibility of differential patterns of correlations for different adjectives certainly merits investigation, since both DeSoto et al. (1965) and Shaver et al. (1974) have suggested that different adjective pairs may be processed in qualitatively different ways.

The correlations with individual sessions are of interest as a further test of the strategy-change hypotheses. According to the spatial-to-linguistic strategy-change hypothesis, one might expect higher correlations with spatial tests in earlier sessions, followed by higher correlations with linguistic tests in later sessions. The linguistic-to-spatial strategy change hypothesis might lead one to make exactly the opposite prediction.

Looking first at the combined data, we see that all correlations are statistically significant except those with the verbal composite in Experiments 1 and 2. This pattern of results can be understood in terms of the mixed model.

According to this model, there is a key difference in strategy between subjects in Experiments 1 and 2 and subjects in Experiments 3 and 4, namely, a reduced emphasis in the latter experiments upon spatial seriation accompanied by checking of previous linguistic encodings and possible use of the linguistic noncongruence operation. These changes in strategy should result in an increase in the relative contribution of verbal ability to the solution of linear syllogisms in Experiments 3 and 4, and possibly a decrease in the relative contribution of spatial-abstract ability. The correlations show a pronounced increase in the verbal contribution, and a possible decrease in the spatial-abstract contribution.

The patterns of correlations for the individual adjectives do not show any consistent trends across experiments. Although there are trends that might be viewed as suggestive in the context of single experiments, these trends do not hold up when considered in the context of the entire set of data. These correlations, like the model fits for individual adjectives, suggest that a single model is likely to account for processing strategy for each of the three adjective pairs.

The correlations for individual sessions tell much the same story. Although there are isolated patterns within single experiments, no trend seems to hold up when the experiments are considered in conjunction. In particular, there is no suggestion in the data that subjects rely upon either a spatial or linguistic strategy in earlier sessions, and then switch to the other strategy in later sessions. These data, like the modeling data, suggest that a single model is likely to account for the data in every session.

To summarize, the correlational data are consistent with the modeling data in suggesting that a single model can account for performance across both adjectives and sessions. The data reviewed so far favor the mixed model as this single model.

COMPARISON OF THE MODELS ON PREVIOUSLY PUBLISHED DATA

Preview

The results presented in the previous part of the article were generally supportive of the mixed model in the context of the present set of experiments. How does the mixed model compare to the alternative models, however, in its ability to account for previously published results? This question is addressed in the present part of the article. First, qualitative aspects of model fits will be discussed, and then quantitative ones.

Qualitative Aspects of Fit

Representation of Marked versus Unmarked Adjectives in Memory

Potts and Scholz (1975) reported two findings that led them to believe that marked and unmarked adjectives are represented in the same form in memory, regardless of the way in which premises are stated. The first finding was that "when subjects are given sufficient time to study the premises prior to answering the question, reaction time to the question 'Who is best?' is shorter than reaction time to the question 'Who is worst"'" (p. 445). All of the models as formulated in this article can handle this finding. The finding is consistent with the notion that marked adjectives take longer to encode (whether the encoding is linguistic, spatial, or both) than do unmarked ones. One would therefore expect longer solution times as a function of longer times spent in encoding the marked adjective in the question "Who is worst?" Potts and Scholz recognized the differential encoding interpretation as an alternative to their own interpretation of their finding as indicating a single form of storage.

Potts and Scholz's second finding was that there is no effect of noncongruence in a separate-stages (precueing) paradigm. This finding is consistent only with the mixed model, which asserts that subjects do not check for noncongruence of the question with their linguistic encoding of the answer when they are encouraged, as they were in Potts and Scholz's separate-stages paradigm, to form a sharply defined spatial array.

Ability of Subjects to Answer Unexpected Questions as a Function of Practice

Wood, Shotter, and Godden (1974) found that with increasing practice in solving five-term series problems, subjects showed "a general reduction in the ability to answer unexpected questions based on the information just utilized" (p. 255). The authors interpreted these findings as corroborating "the claim that subjects who are naive, with respect to series problems, generally tend to adopt a representational strategy while those who are more experienced tend to develop a nonrepresentational one" (p. 255). An alternative explanation, which has nothing to do with alternative modes of problem representation, is that with increasing amounts of practice, subjects establish a set for solving the problems at hand. The more problems of a similar nature the subjects are given to solve, the more likely they are to fail to solve a set-breaker. This set or functional fixedness effect is a common one in problem-solving tasks (see, for example, Duncker, 1945; Luchins, 1942), and seems applicable here. Although there was a control group in the Wood et al. experiment, the nature of the task given to the control group was such that any set that might have built up was irrelevant to the unexpected question, and hence would not have been expected to interfere with the subjects' answering it.

Difficulties People Have in Answering the Question "Where is It?"

Clark (1972a) performed a series of experiments in which subjects were instructed to insert an object into a visually presented array. The experiment most relevant to the present discussion is the third. In this experiment, one group would be presented with "32 displays each constructed from one of eight sentences--*Blue is higher than* (*is lower than*, *isn't as high as*, *isn't as low as*) *pink*, plus the same four sentences with *blue* and *pink* interchanged--and from one of four different pairs of colored lines--black on top and blue on bottom, blue on top and black on bottom, black on top and pink on bottom, and pink on top and black on bottom....The Ss were told to indicate whether the missing blue or pink line went above or below both of the lines on the right by pressing the top or bottom button on their response panel" (p. 271). A second group of subjects received identical problems, except that the terms better and worse were substituted for higher and lower. For each group, half of the items were affirmative and half were negative equative; further, half of the items had determinate answers and half did not (so that subjects could not tell where the missing line went, and had to indicate as much).

The placement task used by Clark bears certain structural similarities to the linear syllogisms task, and Clark (1972a) compared data from this task to data from items alleged to be structurally analogous in Clark's (1969a) linear syllogisms data. The data from the two tasks showed qualitatively different patterns, and the correlations between latencies in the placement task and errors in the linear syllogisms task was only .55 for determinate items (the only type considered in this article). One is therefore obliged to conclude that the placement task bore only a weak relation to the linear syllogisms task.

This conclusion presents a problem for a model of linear syllogistic reasoning only if (a) one claims that there is an isomorphism between certain physical placement tasks and linear syllogisms tasks, and (b) one accepts Clark's claim that his placement task is one for which there should be an isomorphism, if there should be an isomorphism for any such task. Proponents of the spatial and mixed models would probably accept the first claim and reject the second claim almost unanimously (see, for example, Huttenlocher, 1968; Huttenlocher & Higgins, 1972). On the one hand, theorists positing the use of spatial imagery in linear syllogistic reasoning seem to agree that internal spatial arrays are analogous at some level to external physical arrays that are viewed in everyday life. On the other hand, none of these theorists would argue that any physical arrangement task that is isomorphic or nearly isomorphic to the linear syllogisms task should result in the same structures and processes as are used in the linear syllogisms task. Indeed, such a claim would be foolish in light of results such as those of Hayes and Simon (1977), which show that even carefully controlled problem isomorphs can lead to vastly different representations and processes if they are presented with the appropriate surface structural differences. As both these authors and Huttenlocher and Higgins (1972) point out, representations and processes are highly sensitive to surface structural differences; and an abundance of such differences exist between Clark's (1972a) placement task and the linear syllogisms task. Huttenlocher and her associates have found a number of placement tasks that do seem to yield results paralleling those from linear syllogisms tasks (see Huttenlocher & Higgins, 1971, for a review); Clark (1972a) has found a class of placement tasks that does not yield results paralleling those from linear syllogisms tasks. The precise conditions under which paralellism does or does not result remain to be specified.

Quantitative Aspects of Fit

Although a number of data sets have been reported in the literature, most of them do not contain even the minimum range of item types that would permit the three models to be distinguished (for example, DeSoto et al., 1965; Handel et al., 1968; Huttenlocher, 1968). Thus, the number of data sets that could be used for quantitative comparison was very limited.

Adult Subjects

Clark (1969b). Clark has published geometric mean latencies for the 32 uncued item types used in the present experiments. The quality of the data are suspect, however, as even Clark (1971) implies. The latencies are based upon only 13 subjects, with just three observations per subject. Moreover, Clark threw out the longest latency for each subject for each item (33% of the observations), and also all error responses (7% of the observations). Values of R^2 for the mixed, linguistic, and spatial models respectively were .63, .72, and .53. These results thus favor the linguistic model.

Clark (1969a). Modeling could again be done on data from the 32 uncued item types used in the present experiments. In this experiment, Clark gave subjects 10 sec to solve each problem. An error was counted if the subject either responded incorrectly or failed to respond at all in the 10 sec. The present modeling is of the proportion of errors for each problem type. Modeling of the logarithm of the number of correct responses yielded comparable results. Values of R^2 for the mixed, linguistic, and spatial models were .59, .65, and .60. These data thus give a slight edge to the linguistic model.

Potts and Scholz (1975). The eight data points from Experiment 1, Group 1, of Potts and Scholz (1975) also provided an adequate basis for distinguishing among models. The values of R^2 were .86, .73, and .48 for the mixed, linguistic, and spatial models respectively. The data thus support the mixed model.

Child Subjects

Keating and Caramazza (1975). These authors used the same 10 sec deadline procedure as did Clark (1969a). Their subjects were bright and average fifth and seventh grade children. Their data permitted modeling of error rates for eight item types. The respective values of R^2 for the mixed, linguistic, and spatial models were .96, .71, and .68 for average fifth graders, .84, .99, and .83 for bright fifth graders, .70, .96, and .70 for average seventh graders, and .52, .94, and .52 for bright seventh graders. For the combined fifth graders, values of R^2 were .92, .86, and .77 for the three models; for the combined seventh graders, values of R^2 were .60, .96, and .60. For the combined average students, values of R^2 were .88, .86, and .75; for the combined bright students, values of R^2 were .70, .98, and .70. Finally, for the total sample, values of R^2 were .81, .93, and .76.

Although the replicability of these data obviously needs to be established, the data are of particular interest in suggesting a developmental shift in error patterns: The shift is between the mixed model and the linguistic model, with use of the linguistic model associated with greater age and brightness.

Hunter (1957). Hunter tested 11- and 16-year olds on linear syllogisms using the relations *happier-sadder* and *taller-shorter*. His article contains latencies that can be modeled for eight distinct data points. The respective values of R^2 for the mixed, linguistic, and spatial models were .75, .74, and .82 for the 11-year olds, .66, .36, and .53 for the 16-year olds, and .75, .68, and .73 for the combined age groups. Some of the unconstrained parameter estimates for the linguistic model were negative, and so these were forced to be nonnegative in the final linear modeling. These data, like Keating and Caramazza's, suggest the possibility of a developmental trend, but here it is from the spatial to the mixed model. The linguistic model never performed best.

Conclusion

The data from previous research are confusing and contradictory. The results of Clark (1969b) and of the two developmental studies must be interpreted with caution, the former because of the massive deletion of observations, the latter because of the availability of only eight data points for modeling. The Clark (1969a) data appear to be reliable, however, and on their face contradict the mixed model. The reason for this contradiction has now been discovered, and is discussed in detail elsewhere (Sternberg, Note 3). In essence, the contradiction arises because of the modeling of error (or similarly, log correct) data rather than latency data.

GENERAL DISCUSSION

Near the beginning of the article, three important theoretical questions were posed regarding representation and process in transitive inference. The time has now come to see how these questions can be answered on the basis of the theory and data presented in this article.

Representation of Information

The evidence presented in this article suggests strongly that both linguistic and spatial representations for information are used during the course of solution of transitive inference problems. Subjects first decode the linguistic surface structure of the premises into a linguistic deep structure, and then recode the linguistic deep structure into a spatial array. Both the linguistic deep structure and the spatial array are available for search and retrieval processes that occur after recoding has taken place. Noncongruence, when used, operates upon the linguistic representation, while response search operates upon the spatial representation.

Processing of Information

The preferred mixed model accounts for transitive inference in the solution of linear syllogisms in terms of 12 elementary information-processing components, not all of which are used in every type of problem and not all of which have been estimated as separate parameters in the preceding experiments. Of the 12 processes, six were hypothesized to be linguistic (premise reading, linguistic encoding of unmarked adjectives, linguistic encoding of marked adjectives, noncongruence, question reading, negation), five were hypothesized to be spatial (seriation, spatial encoding of unmarked adjectives, spatial encoding of marked adjectives, pivot search, response search), and one was hypothesized to be neutral (response). Negation turned out to be spatial. The operations were found to differ widely in their latencies and in their contributions to individual differences in overall performance.

Generality of Representations and Processes

Across modes of problem presentation. Linear syllogisms were presented (a) with and without precueing, (b) with the question first and with the question last, (c) with adjective pairs differing within and between subjects. Regardless of the mode of presentation, the mixed model was found to perform substantially better than any of the alternative models.

Across adjectives. Four different adjective pairs--taller-shorter, older-younger, better-worse, faster-slower--were used in the course of the four experiments. Essentially the same results were obtained with each. These results seem to support the generality of the representations and processes of the mixed model across adjectives. Of course, this generality is consistent with the earlier findings of DeSoto et al. (1965) and of Handel et al. (1968) that subjects may differ in the directions they use between and within adjective pairs for representing spatial arrays.

Across sessions. The numbers of sessions in the four experiments ranged from one to three. The evidence supported the generality of the representations and processes of the mixed model across all sessions. There was no evidence of the kinds of strategy shifts suggested either by Wood et al. (1974) and Wason and Johnson-Laird (1972) or by Shaver et al. (1974). As would be expected, there was evidence that subjects speed up with increasing practice.

Across subjects. Analyses of individual data revealed a striking consistency in the superiority of the mixed model. The mixed model was not used universally, however. About 13% of the subjects in the four experiments showed evidence of using a strategy more closely approximated by either the linguistic or the spatial model. Moreover, results from previous investigators suggest that there may be population differences in strategies, especially across ages.

Across tasks. The present experiments have tested the generalizability of the representations and processes of the mixed model across a variety of experimental paradigms using linear syllogisms. The presence or absence of a particular operation in the linear syllogisms task obviously does not guarantee the presence or absence of that operation in other tasks. How generalizable are the operations identified in the mixed model? This question can be answered in two different ways.

First, at least some generalizability has already been shown by the significant correlations of each component latency with at least one of the three reference-ability composites. The correlations show that the patterns of individual differences generated by the component processes are not specific to the linear syllogisms task, but are common to reference tests that have been shown to measure abilities called upon in a wide variety of psychometric and other tests. In particular, the correlation of each component latency with the type of ability it is hypothesized to represent demonstrates the convergent validity of the information-processing component. The lack of correlation of each component latency with a type of ability it is hypothesized not to represent demonstrates the discriminant validity of the information-processing component. Only two mispredictions arose. Negation showed a clear convergent-discriminant pattern, but it was spatial rather than linguistic. In addition to strong correlations with the spatial-abstract tests, response search also showed weak but significant correlations with the verbal tests, and these latter correlations were interpreted in terms of individual differences in times for reading names from the spatial arrays.

Second, the generalizability of encoding, negation, marking, and noncongruence operations to a wide variety of comprehension and reasoning tasks has been amply demonstrated in past research. (See Carpenter & Just, 1975; Clark, 1973; Clark & Chase, 1972; Trabasso, 1972; for a comprehensive review of relevant literature.) It should be noted that although this past research has repeatedly identified these information-processing components as contributors to solution latency in a variety of tasks, the research has not adequately distinguished which of the identified operations are linguistic and which are spatial.

Consider, for example, one of the most widely studied comprehension tasks, the sentence-picture comparison task. A subject is shown a sentence and a picture that either may or may not illustrate the situation described by the sentence. The subject must indicate whether or not the sentence describes the situation depicted in the picture. For example, a typical sentence might be "Star is above plus," with $\overset{+}{*}$ as the accompanying picture. In order to solve problems such as these, the subject must (a) encode the sentence and picture, (b) comprehend the negation if one appears, (c) spend additional time comprehending the marked description (*below*) if one appears, (d) spend extra time setting a "truth index" to *false* if the embedding strings of sentence and picture are noncongruent, or if the embedded strings of sentence and picture are noncongruent in the deep-structural propositional representations into which the surface structures have been encoded, (e) respond. Although Clark and Chase (1972) view each of these operations as linguistic, the operations may be viewed as consistent with either the spatial or mixed model as well, with each operation viewed in the same way it is viewed for the solution of linear syllogisms.

The pivot search and response search operations as formulated in this article have not appeared previously in the literature, and so their generalizability has yet to be demonstrated fully. However, there is some evidence that provides further support for the plausibility of these operations. If pivot search as formulated by the mixed model is used, then negative equative problems should show end-anchoring effects for their linguistically converted (recoded) form. Huttenlocher et al. (1970) have found such a result in their task requiring manipulation of physical objects rather than just abstract terms. Findings such as these and those of Shepard and Metzler (1971) and Kosslyn (1975) suggest that in scanning visualized arrays, subjects proceed in much the same way they do in scanning physical arrays.

Present research is directed toward further demonstrations of the generalizability of the components of the mixed model. In one ongoing study, the mixed model is being extended to and tested on series problems with from two to six terms. In a second study, the model is being extended to linear syllogisms with indeterminate solutions. In a third study, the model is being tested on the performance of children from ages of 8 to 16 in solving linear syllogisms. With these and other extensions, it is believed that the mixed model, or an augmented version of it, can be shown to provide some insight into a variety of forms of human cognition.

Appendix

Table A shows predicted values for the mixed model and observed values of latencies for each of the 32 uncued data points in each of the four experiments.

Insert Table A about here

Reference Notes

1. Anderson, J. R. *The status of arguments concerning representations for mental imagery*.
2. Holyoak, K. J. *Symbolic processes in mental comparisons*. Unpublished doctoral thesis, Stanford University, 1976.
3. Sternberg, R. J. *A proposed resolution of curious conflicts in the literature on linear syllogisms*. Manuscript submitted for publication, 1977.
4. Wood, D. J. *The nature and development of problem-solving strategies*. Unpublished doctoral thesis, University of Nottingham, 1969.
5. Clark, H. H., personal communication, August, 1977.

References

Bower, G. H. Adaptation-level coding of stimuli and serial position effects. In M. H. Appley (Ed.), *Adaptation-level theory*. New York: Academic Press, 1971.

Bryant, P. E., & Trabasso, T. Transitive inferences and memory in young children. *Nature*, 1971, *232*, 456-458.

Burt, C. The development of reasoning in school children. *Journal of Experimental Pedagogy*, 1919, *5*, 68-77.

Carpenter, P. A., & Just, M. A. Sentence comprehension: A psycholinguistic processing model of verification. *Psychological Review*, 1975, *82*, 45-73.

Chomsky, N. *Aspects of the theory of syntax*. Cambridge, Mass.: MIT Press, 1965.

Clark, H. H. The influence of language in solving three-term series problems. *Journal of Experimental Psychology*, 1969, *82*, 205-215. (a)

Clark, H. H. Linguistic processes in deductive reasoning. *Psychological Review*, 1969, *76*, 387-404. (b)

Clark, H. H. More about "Adjectives, comparatives, and syllogisms": A reply to Huttenlocher and Higgins. *Psychological Review*, 1971, *78*, 505-514.

Clark, H. H. Difficulties people have answering the question "Where is it?" *Journal of Verbal Learning and Verbal Behavior*, 1972, *11*, 265-277. (a)

Clark, H. H. On the evidence concerning J. Huttenlocher and E. T. Higgins' theory of reasoning: A second reply. *Psychological Review*, 1972, *79*, 428-432.

Clark, H. H. Semantics and comprehension. In T. A. Sebeok (Ed.), *Current trends in linguistics*, Vol. 12: *Linguistics and adjacent arts and sciences*. The Hague: Mouton, 1973. Pp. 1291-1337.

Clark, H. H., & Chase, W. On the process of comparing sentences against pictures. *Cognitive Psychology*, 1972, *3*, 472-517.

Cooper, L. A., & Shepard, R. N. Chronometric studies of the rotation of mental images. In W. Chase, Visual information processing. New York: Academic Press, 1973. Pp. 75-176.

DeSoto, C. B., London, M., & Handel, S. Social reasoning and spatial paralogic. Journal of Personality and Social Psychology, 1965, 2, 513-521.

Donaldson, M. A study of children's thinking. London: Tavistock, 1963.

Duncker, K. On problem-solving. Psychological Monographs: General and Applied, 1945, 58(5, Whole No. 270).

French, J. W., Ekstrom, R. B., & Price, I. A. Kit of reference tests for cognitive factors. Princeton: Educational Testing Service, 1963.

Handel, S., DeSoto, C. B., & London, M. Reasoning and spatial representations. Journal of Verbal Learning and Verbal Behavior, 1968, 7, 351-357.

Hayes, J. R., & Simon, H. A. Psychological differences among problem isomorphs. In N. J. Castellan, Jr., D. B. Pisoni, & G. R. Potts (Eds.), Cognitive Theory (Vol. 2). Hillsdale, N.J.: Lawrence Erlbaum Associates, 1977.

Hunter, I. M. L. The solving of three term series problems. British Journal of Psychology, 1957, 48, 286-298.

Huttenlocher, J. Constructing spatial images: A strategy in reasoning. Psychological Review, 1968, 75, 550-560.

Huttenlocher, J., Eisenberg, K., & Strauss, S. Comprehension: Relation between perceived actor and logical subject. Journal of Verbal Learning and Verbal Behavior, 1970, 9, 334-341.

Huttenlocher, J., & Higgins, E. T. Adjectives, comparatives, and syllogisms. Psychological Review, 1971, 78, 487-504.

Huttenlocher, J., & Higgins, E. T. On reasoning, congruence, and other matters. Psychological Review, 1972, 79, 420-427.

Huttenlocher, J., Higgins, E. T., Milligan, C., & Kauffman, B. The mystery of the "negative equative" construction. Journal of Verbal Learning and Verbal Behavior, 1970, 9, 334-341.

Huttenlocher, J., & Strauss, S. Comprehension and a statement's relation to the situation it describes. Journal of Verbal Learning and Verbal Behavior, 1968, 7, 300-304.

Johnson-Laird, P. N. The three-term series problem. Cognition, 1972, 1, 57-82.

Keating, D. P., & Caramazza, A. Effects of age and ability on syllogistic reasoning in early adolescence. Developmental Psychology, 1975, 11, 837-842.

Kosslyn, S. M. Information representation in visual images. Cognitive Psychology, 1975, 7, 341-370.

Kosslyn, S. M., Pomerantz, J. R. Imagery, propositions, and the form of internal representations. Cognitive Psychology, 1977, 9, 52-76.

Lawson, R. Representation of individual sentences and holistic ideas. Journal of Experimental Psychology: Human Learning and Memory, 1977, 3, 1-9.

Luchins, A. S. Mechanization in problem solving. Psychological Monographs, 1942, 54(Whole No. 248).

Lutkus, A. D., & Trabasso, T. Transitive inferences in preoperational retarded adolescents. American Journal of Mental Deficiency, 1974, 78, 599-606.

Moyer, R. S. Comparing objects in memory: Evidence suggesting an internal psychophysics. Perception and psychophysics, 1973, 13, 180-184.

Moyer, R. S., & Bayer, R. H. Mental comparison and the symbolic distance effect. Cognitive Psychology, 1976, 8, 228-246.

Nisbett, R. E., & Wilson, T. D. Telling more than we can know: Verbal reports on mental processes. Psychological Review, 1977, 84, 231-259.

Piaget, J. Une forme verbal de la comparison chez l'enfant. Archives de Psychologie, 1921, 141-172.

Piaget, J. *Judgment and reasoning in the child*. London: Routledge & Kegan Paul, 1928.

Piaget, J. *The language and thought of the child*. New York: Meridian Books Inc., 1955.

Piaget, J. *Genetic epistemology*. Translated by E. Duckworth. New York: Columbia University Press, 1970.

Potts, G. R. Information processing strategies used in the encoding of linear orderings. *Journal of Verbal Learning and Verbal Behavior*, 1972, *11*, 727-740.

Potts, G. R. Storing and retrieving information about ordered relationships. *Journal of Experimental Psychology*, 1974, *103*, 431-439.

Potts, G. R., & Scholz, K. W. The internal representation of a three-term series problem. *Journal of Verbal Learning and Verbal Behavior*, 1975, *14*, 439-452.

Pylyshyn, Z. W. What the mind's eye tells the mind's brain: A critique of mental imagery. *Psychological Bulletin*, 1973, *80*, 1-24.

Riley, C. A., & Trabasso, T. Comparatives, logical structures and encoding in a transitive inference task. *Journal of Experimental Child Psychology*, 1974, *17*, 187-203.

Shaver, P., Pierson, L., & Lang, S. Converging evidence for the functional significance of imagery in problem solving. *Cognition*, 1974, *3*, 359-375.

Shepard, R. N., & Metzler, J. Mental rotation of three-dimensional objects. *Science*, 1971, *171*, 701-703.

Sternberg, R. J. Component processes in analogical reasoning. *Psychological Review*, 1977, *84*, 353-378.

Sternberg, R. J. *Intelligence, information processing, and analogical reasoning: The componential analysis of human abilities*. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1977. (b)

Trabasso, T. Mental operations in language comprehension. In J. B. Carroll & R. O. Freedle (Eds.), *Language comprehension and the acquisition of knowledge*. Washington, D.C.: Winston, 1972. Pp. 113-137.

Trabasso, T. Representation, memory, and reasoning: How do we make transitive inferences? In A. D. Pick (Ed.), *Minnesota symposia on child psychology* (Vol. 9). Minneapolis: University of Minnesota Press, 1975. Pp. 135-172.

Trabasso, T., & Riley, C. A. On the construction and use of representations involving linear order. In R. L. Solso (Ed.), *Information processing and cognition: The Loyola Symposium*. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1975.

Trabasso, T., Riley, C. A., & Wilson, E. G. The representation of linear order and spatial strategies in reasoning: A developmental study. In R. Falmagne (Ed.), *Reasoning: Representation and process*. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1975.

Wason, P. C., & Johnson-Laird, P. N. *Psychology of reasoning: Structure and content*. London: B. T. Batsford Ltd., 1972.

Wood, D., Shotter, J., & Godden, D. An investigation of the relationships between problem solving strategies, representation and memory. *Quarterly Journal of Experimental Psychology*, 1974, *26*, 252-257.

Footnotes

Portions of this article were presented at the Mathematical Psychology Meetings in New York, September, 1976, and at the Psychonomic Society Meetings in St. Louis, November, 1976. I am grateful to Judy Tschirgi for experimental assistance in and valuable suggestions regarding Experiment 1. Barbara Conway tested subjects in Experiments 2 and 3; Elizabeth Hartka tested subjects in Experiment 4. Experiment 4 arose out of conversations with Alice Healy and Michael Kubovy. Experiment 1 was supported by NIMH Grant MH-13950-07 to Gordon Bower and a National Science Foundation predoctoral fellowship to the author. Experiments 2, 3, and 4, and preparation of this article, were supported by NSF Grant BNS-76-05311 to Robert Sternberg. Requests for reprints should be sent to Robert J. Sternberg, Department of Psychology, Box 11A Yale Station, Yale University, New Haven, Connecticut 06520.

[1] Solution latencies for practice items were not recorded, and of course, it is not possible to determine what strategy or strategies were used during the solution of these items.

Table 1

Estimated Parameters of Mathematical Models

Parameter	Operations	Experiment 1	2	3-4
	Mixed Model			
ENC+$_1$	SER + (2)PR + NMAR1 + NMAR2	x	x	
ENC+$_2$	SER + PR + (.5)NMAR1 + (.5)NMAR2			x
NEG	NEG	x	x	x
MARK	MARK1 + MARK2 - NMAR1 - NMAR2	x	x	x
PSM	PSM	x	x	x
RS	RS	x	x	x
NCON	NCON[a]			x
QR+	QR + (.5)NMAR1 + (.5)NMAR2		x	
RES+$_1$	RES + QR + (.5)NMAR1 + (.5)NMAR2	x		
RES	RES		x	
RES+$_2$	RES + QR + PR + NMAR1 + NMAR2			x
RES++$_1$[b]	RES + QR + SER + (2)PR + (1.5)NMAR1 + (1.5)NMAR2	x	x	x
	Linguistic Model			
ENC+$_3$	(2)PR + NMAR	x	x	
ENC+$_4$	PR + (.5)NMAR			x
NEG	NEG	x	x	x
MARK	MART - NMAR	x	x	x
PSL	PSL	x	x	x
NCON	NCON	x	x	x
QR+	QR + (.5)NMAR		x	
RES+$_4$	RES + QR + (.5)NMAR + SR	x		
RES+$_5$	RES + SR		x	

Table 1 (Continued)

Parameter	Operations	Experiment 1	2	3-4
RES+$_6$	RES + QR + PR + NMAR + SR			x
RES++$_2$^{b}	RES + QR + (2)PR + (1.5)NMAR	x	x	x
	Spatial Model			
ENC+$_5$	SER + (2)PR + NMAR	x	x	
ENC+$_6$	SER + PR + (.5)NMAR			x
NEG	NEG	x	x	x
MARK	MART - NMAR	x	x	x
PSS	PSS	x	x	x
SERN	SERN	x	x	x
QR+	QR + (.5)NMAR		x	
RES+$_1$	RES + QR + (.5)NMAR	x		
RES	RES		x	
RES+$_2$	RES + QR + PR + NMAR			x
RES++$_1$^{b}	RES + QR + SER + (2)PR + (1.5)NMAR	x	x	x

[a] Optional

[b] Uncued items only

Table 2

Three-Term Series Problem Types

Values of Independent Variables Multiplying Parameters

Item No.	Premises	Question	Common			Mixed		Linguistic		Spatial	
			EMC+	MARK	NEG	PSM	RS	NCON	PSL	PSS	SERN
			Uncued Items								
1	A>B B>C	>	1	0	0	0	1	0	1	1	0
2	A>B B>C	<	1	1	0	0	0	1	1	1	0
3	B>C A>B	>	1	0	0	0	0	0	0	1	1
4	B>C A>B	<	1	1	0	0	1	1	0	1	1
5	C<B B<A	>	1	2	0	0	0	1	1	1	1
6	C<B B<A	<	1	3	0	0	1	0	1	1	1
7	B<A C<B	>	1	2	0	0	1	1	0	1	0
8	B<A C<B	<	1	3	0	0	0	0	0	1	0
9	A>B C<B	>	1	1	0	0	1	0	0	0	0
10	A>B C<B	<	1	2	0	0	0	0	0	0	0
11	C<B A>B	>	1	1	0	0	0	0	0	0	1
12	C<B A>B	<	1	2	0	0	1	0	0	0	1
13	B<A B>C	>	1	1	0	0	1	1	0	2	0
14	B<A B>C	<	1	2	0	0	0	1	0	2	0
15	B>C B<A	>	1	1	0	0	0	1	0	2	1
16	B>C B<A	<	1	2	0	0	1	1	0	2	1

Table 2 Continued

Three-Term Series Problem Types

Item No.	Premises	Question	Common			Mixed		Linguistic		Spatial	
			ENC+	MARK	NEG	PSM	RS	NCON[a]	PSL	PSS	SERN
17	A≮B B≮C	>	1	2	2	0	1	1	0	1	0
18	A≮B B≮C	<	1	3	2	0	0	0	0	1	0
19	B≮C A≮B	>	1	2	2	1	0	1	1	1	1
20	B≮C A≮B	<	1	3	2	1	1	0	1	1	1
21	C≯B B≯A	>	1	0	2	0	0	0	0	1	1
22	C≯B B≯A	<	1	1	2	0	1	1	0	1	1
23	B≯A C≯B	>	1	0	2	1	1	0	1	1	0
24	B≯A C≯B	<	1	1	2	1	0	1	1	1	0
25	A≮B C≯B	>	1	1	2	1	1	1	0	2	0
26	A≮B C≯B	<	1	2	2	1	0	1	0	2	0
27	C≯B A≮B	>	1	1	2	1	0	1	0	2	1
28	C≯B A≮B	<	1	2	2	1	1	1	0	2	1
29	B≯A B≮C	>	1	1	2	0	1	0	0	0	0
30	B≯A B≮C	<	1	2	2	0	0	0	0	0	0
31	B≮C B≯A	>	1	1	2	0	0	0	0	0	1
32	B≮C B≯A	<	1	2	2	0	1	0	0	0	1

[a] Optional in mixed model

Table 3

Two-Term Series Problem Types

Values of Independent Variables Multiplying Parameters

Item No.	Premise	Question	Common			Mixed		Linguistic		Spatial	
			ENC+	MARK	NEG	PSM	RS	NCON	PSL	PSS	SERN
1	A>B	>	0	0	0	0	0	0	0	0	0
2	A>B	<	0	1	0	0	0	1	0	0	0
3	B<A	>	0	1	0	0	0	1	0	0	0
4	B<A	<	0	2	0	0	0	0	0	0	0
5	B≯A	>	0	0	1	0	0	0	0	0	0
6	B≯A	<	0	1	1	0	0	1	0	0	0
7	A≮B	>	0	1	1	0	0	1	0	0	0
8	A≮B	<	0	2	1	0	0	0	0	0	0

Table 4

Basic Statistics for Data Used in Modeling

	Experiment 1		Experiment 2		Experiment 3		Experiment 4	
	$\bar{X}$	$S_{\bar{X}}$	$\bar{X}$	$S_{\bar{X}}$	$\bar{X}$	$S_{\bar{X}}$	$\bar{X}$	$S_{\bar{X}}$
Uncued Data								
Overall	7285	177	7489	188	7002	170	7069	161
Adjective Subsets								
Tall-Short	7597	238	8000	213	7046	170	6602	147
Old-Young	6919	139	----	---	----	---	----	---
Good-Bad	----	---	7712	223	7022	181	7096	166
Fast-Slow	7340	247	6753	206	6937	189	7509	195
Session Subsets								
Session 1	7551	248	7809	238	7976	233	7069	161
Session 2	7152	167	7137	186	7022	174	----	---
Session 3	----	---	7518	228	6021	134	----	---
Combined Uncued and Cued Data								
Overall	4475	365	5072	321	5093	326	6266	288
Precueing[a]								
Overall	7077	182	7098	202	3185	134	3053	114

Note: All response times are expressed in milliseconds.

[a] Precueing consists of two-term series problems in Experiments 3 and 4. Two- and three-term series problems in these experiments were presented in separate trials.

Table 5

Performance of Models in Predicting Latency Data

		Proportion of Variance Accounted For (R^2)			
Data Set	**Model**	**Experiment 1**	**Experiment 2**	**Experiment 3**	**Experiment 4**
		Uncued Items only			
Overall	Mixed	.813	.740	.844	.884
	Linguistic	.600	.592	.689	.644
	Spatial	.568	.587	.577	.576
Tall-Short	Mixed	.497	.565	.748	.714
	Linguistic	.377	.534	.707	.533
	Spatial	.415	.499	.529	.499
Old-Young	Mixed	.593	----	----	----
	Linguistic	.346	----	----	----
	Spatial	.302	----	----	----
Good-Bad	Mixed	----	.588	.742	.822
	Linguistic	----	.446	.629	.584
	Spatial	----	.484	.599	.532
Fast-Slow	Mixed	.757	.641	.834	.864
	Linguistic	.638	.451	.564	.632
	Spatial	.589	.495	.461	.552
Session 1	Mixed	.551	.646	.709	.884
	Linguistic	.379	.520	.633	.644
	Spatial	.379	.490	.491	.576

Table 5 Continued

		Proportion of Variance Accounted For (R^2)			
Data Set	Model	Experiment 1	Experiment 2	Experiment 3	Experiment 4
Session 2	Mixed	.795	.540	.812	----
	Linguistic	.616	.433	.669	----
	Spatial	.581	.446	.593	----
Session 3	Mixed	----	.539	.814	----
	Linguistic	----	.435	.544	----
	Spatial	----	.482	.495	----
Combined Uncued and Cued Items					
Overall	Mixed	.985	.971	.974	.971
	Linguistic	.974	.962	.957	.923
	Spatial	.973	.960	.940	.911
Precueing[a]					
Overall	Mixed	.831	.649	----	----
	Linguistic	.663	.619	----	----
	Spatial	.772	.523	----	----

---

[a] Two-term series problems did not contain sufficient constraints to distinguish models.

Table 6

Tests of Residuals

Predicted versus Observed Values for Models of Transitive Inference

Uncued Items

Within Experiment

(Internal Consistency across Item Types)

	Experiment			
Model	1	2	3	4
Mixed	-.01	.31*	.36*	.64***
Linguistic	.36*	.51**	.60***	.85***
Spatial	.39*	.54***	.69***	.87***

Between Experiments

(Consistency across Item Types)

	Experimental Comparison					
Model	1-2	1-3	1-4	2-3	2-4	3-4
Mixed	.28	.36*	.39*	.22	-.03	.51***
Linguistic	.66***	.65***	.58***	.78***	.74***	.81***
Spatial	.63***	.69***	.72***	.69***	.62***	.84***

Note: All tests are one-tailed.

*$p < .05$

**$p < .01$

***$p < .001$

Table 7

Performance of Models in Predicting Individual Latency Data

Model	Uncued Items Only		Uncued and Cued Items	
	R^2	Best[a]	R^2	Best[a]
		Experiment 1		
Mixed	.361	10	.855	9
Linguistic	.316	2	.846	2
Spatial	.308	4	.851	5
		Experiment 2		
Mixed	.340	13	.857	12
Linguistic	.261	2	.845	4
Spatial	.262	3	.842	2
		Experiment 3		
Mixed	.510	14	.870	13
Linguistic	.432	2	.855	3
Spatial	.361	2	.831	2
		Experiment 4		
Mixed	.404	40	.693	43
Linguistic	.298	8	.643	6
Spatial	.268	6	.632	5

[a] Number of cases in which each model best fit the data of individual subjects.

Proportions of Cases in Which One Model Performed Substantially Better than Another Model

Uncued Items

Better Model	Worse Model		
	Mixed	Linguistic	Spatial
Mixed	----- -----	.65 (.49)	.72 (.53)
Linguistic	.08 (.05)	----- -----	.37 (.18)
Spatial	.06 (.03)	.19 (.06)	----- -----

Note: Top number in each cell indicates difference in R^2 of at least .05. Bottom (parenthesized) number in each cell indicates difference in R^2 of at least .10.

Table 9

Mixed Model Parameter Estimates for Different Data Sets

Parameter	Solution Time Estimates (Uncued)				Solution Time Estimates (Uncued & Cued)				Cue Time Estimates	
	Experiment									
	1	2	3	4	1	2	3	4	1	2
ENC+$_1$	----	----	----	----	4648** (130)	4666** (190)	----	----	----	----
ENC+$_2$	----	----	----	----	----	----	2986** (184)	3124** (152)	----	----
NEG	351** (100)	373** (126)	100 (92)	224** (75)	351** (79)	366** (85)	184* (86)	244** (71)	439** (96)	369* (157)
MARK	416** (95)	461** (119)	451** (85)	429** (69)	337** (65)	412** (87)	307** (73)	380** (63)	532** (111)	399** (148)
PSM	1136** (232)	1086** (290)	1256** (222)	1020** (181)	1136** (183)	1045** (196(	1154** (226)	1008** (174)	1179** (223)	1085** (363)
RS	616** (164)	595** (205)	522** (147)	656** (120)	380** (92)	695** (138)	522** (163)	656** (118)	----	757** (256)
NCON	----	----	570** (157)	438** (128)	----	----	538** (119)	396** (111)	----	----

Table 9 Continued

Parameter	Solution Time Estimates (Uncued)				Solution Time Estimates (Uncued & Cued)				Cue Time Estimates	
	Experiment									
	1	2	3	4	1	2	3	4	1	2
QR+	----	----	----	----	----	393** (145)	----	----	----	----
$\text{RES+}_1{}^a$	----	----	----	----	1307	----	----	----	----	----
RES^a	----	----	----	----	----	836	----	----	----	----
$\text{RES+}_2{}^a$	----	----	----	----	----	----	2517	2353	----	----
RES++^a	5718	5855	5365	5399	----	----	----	----	----	----

Note: All parameter estimates are in milliseconds. Standard errors of the parameter estimates (also expressed in milliseconds) are shown in parentheses directly beneath the parameter estimates.

[a] Estimated as a regression constant, and hence has no associated significance value or standard error

*$p < .05$
**$p < .01$

Table 10

Parameter Estimates for Linguistic and Spatial Models

Uncued Items Only

	Experiment 1	Experiment 2	Experiment 3	Experiment 4
		Linguistic Model		
Negation (NEG)	635** (120)	644** (129)	416** (112)	479** (103)
Marking (MARK)	416** (138)	461** (149)	451** (118)	429** (119)
Noncongruence (NCON)	259 (240)	545* (257)	884** (204)	693** (206)
Linguistic Pivot Search (PSL)	418 (277)	10 (297)	735** (235)	394 (237)
Response[a] (RES_{++2})	5793	5878	5285	5501
		Spatial Model		
Negation (NEG)	635** (125)	644** (129)	414** (119)	479** (112)
Marking (MARK)	416** (144)	461** (150)	451** (137)	429** (130)
Spatial Pivot Search (PSS)	154 (176)	347 (183)	618** (168)	433* (159)
Seriation in Nonpreferred Direction ([illegible])	162 (249)	178 (259)	139 (238)	17 (225)

Table 10 Continued

	Experiment 1	Experiment 2	Experiment 3	Experiment 4
Response[a] ($RES{+}{+}_1$)	5791	5716	5223	5505

Note: All parameter estimates are in milliseconds. Standard errors of the parameter estimates (also expressed in milliseconds) are shown in parentheses directly beneath the parameter estimates.

[a] Estimated as a regression constant, and hence has no associated significance value or standard error.

$^{*}p < .05$

$^{**}p < .01$

Table 11

Weights of Parameters in Predicting Individual Differences in Latencies for Solution of Linear Syllogisms

Parameter		Experiment 1	Experiment 2	Experiment 3	Experiment 4
		Standardized Regression Coefficients			
Encoding	(ENC+)	.74	.66	.54	.62
Negation	(NEG)	.44	.20	.17	.20
Marking	(MARK)	.19	.19	.16	.15
Pivot Search	(PSM)	.21	.10	.04	.03
Response Search	(RS)	.07	.14	.14	.16
Question Reading	(QR+)	---	.19	---	---
Noncongruence	(NCON)	---	---	.00	.07
Response	(RES+)	.18	.02	.31	.29

Table 12

Intercorrelations of Parameter Estimates for Individual Subjects

Combined Experiments

	ENC+	NEG	MARK	PSM	RS	QR+	NCON	RES+
ENC+	1.00	.50***	.39***	.58***	.17*	.58**	.28*	.31***
	_1.00_	_.57_***	_.73_***	_.62_***	_.52_***	_.71_*	_.44_*	_.31_***
NEG		1.00	.36***	.37***	.02	-.03	.17	-.04
		_1.00_	_.60_*	_.65_**	_.48_*	_-.10_	_.85_*	_.10_
MARK			1.00	.49***	.44***	.49*	.37**	-.09
			_1.00_	_.79_***	_.56_*	_.93_**	_.49_	_-.05_
PSM				1.00	.32**	.53*	.10	.12
				_1.00_	_.68_***	_.82_*	_.64_	_.22_
RS					1.00	.15	.21	.05
					_1.00_	_.92_**	_.84_*	_.22_
QR+						1.00	----	.18
						_1.00_	----	_.30_
NCON							1.00	.09
							_1.00_	_.21_
RES+								1.00
								_1.00_

Note: Nonitalicized values are for all cases for which parameter could be estimated. Italicized values are for cases for which parameter was statistically significant at .025 level.

*$p < .05$

**$p < .01$

***$p < .001$

Table 13

Correlations between Latencies and Composite Ability Scores

	Verbal	Spatial	Abstract
Solution Latency	-.35***	-.49***	-.52***
	-_.35_***	-_.49_***	-_.52_***
ENC+[a]	-.25**	-.51***	-.58***
	-_.25_**	-_.51_***	-_.58_***
NEG	-.14	-.34**	-.41***
	-_.10_	-_.56_***	-_.40_*
MARK	-.20*	-.36***	-.38***
	-_.26_	-_.65_***	-_.70_***
PSM	-.16	-.25**	-.35***
	-_.18_	-_.38_**	-_.42_***
RS	-.26**	-.35***	-.34***
	-_.28_*	-_.58_***	-_.53_***
NCON	-.31*	-.24	-.22
	-_.41_*	-_.38_	-_.27_
RES+[a]	-.30**	-.09	-.15
	-_.30_**	-_.09_	-_.15_

Note: Nonitalicized values are for all cases for which parameter could be estimated. Italicized values are for cases for which parameter was statistically significant at .025 level.

[a] Partial correlation controlling for participation in Experiments 1 and 2 versus Experiments 3 and 4.

*$p < .05$

**$p < .01$

***$p < .001$

Table 14

Correlations between Composite Solution Latency for

Individual Adjectives and Sessions and Ability Test Composites

	Experiment 1	Experiment 2	Experiment 3	Experiment 4
Combined Data				
Verbal	-.41	-.09	-.46*	-.56***
Spatial	-.74***	-.48*	-.56**	-.31*
Abstract	-.83***	-.54*	-.50*	-.48***
Tall-Short				
Verbal	-.37	.12	-.48*	-.57**
Spatial	-.72***	-.67***	-.58**	-.15
Abstract	-.74***	-.79***	-.55**	-.26
Old-Young				
Verbal	-.38			
Spatial	-.78***			
Abstract	-.79***			
Good-Bad				
Verbal		-.57**	-.53*	-.70***
Spatial		-.48*	-.59**	-.53*
Abstract		-.44*	-.52*	-.68***
Fast-Slow				
Verbal	-.40	.29	-.35	-.70***
Spatial	-.59**	-.45*	-.49*	-.55**
Abstract	-.81***	-.49*	-.40*	-.34
Session 1				
Verbal	-.18	-.13	-.44*	-.56***
Spatial	-.66**	-.48*	-.64**	-.31*
Abstract	-.74***	-.55**	-.48*	-.48***
Session 2				
Verbal	-.49*	-.03	-.39	
Spatial	-.73***	-.50*	-.45*	
Abstract	-.82***	-.53*	-.50*	
Session 3				
Verbal		-.09	-.49*	
Spatial		-.45*	-.47*	
Abstract		-.51*	-.44*	

*$p < .05$

**$p < .01$

***$p < .001$

Table A

Predicted versus Observed Data Points for Uncued Items

Item No.	Experiment 1		Experiment 2		Experiment 3		Experiment 4	
	Predicted	Observed	Predicted	Observed	Predicted	Observed	Predicted	Observed
1	633	640	645	598	589	602	606	585
2	613	629	632	630	639	674	627	593
3	572	584	585	616	536	550	540	561
4	675	659	691	699	691	646	692	669
5	655	639	678	717	684	768	669	712
6	758	734	783	714	724	801	734	796
7	717	693	737	765	736	710	735	770
8	697	607	724	702	672	604	669	651
9	675	623	691	637	634	600	648	646
10	655	620	678	666	627	640	626	639
11	613	703	632	715	582	592	583	589
12	717	822	737	683	679	625	691	659
13	675	691	691	743	691	694	692	712
14	655	659	678	709	684	667	669	670
15	613	601	632	548	639	582	627	555
16	717	737	737	807	736	785	735	736
17	787	809	812	797	756	776	780	789
18	767	752	798	688	692	647	713	635
19	839	928	861	894	829	861	816	797
20	942	949	967	951	870	872	881	865
21	642	624	660	652	556	580	585	611
22	745	672	766	781	711	679	737	731
23	817	766	828	768	734	748	752	743
24	797	794	815	724	784	716	773	800
25	859	852	874	863	836	837	839	797
26	839	859	861	929	829	873	816	862
27	797	761	815	775	784	778	773	789
28	900	880	920	1037	882	864	882	880
29	745	808	766	814	654	668	693	715
30	725	725	752	807	647	635	671	688
31	684	679	706	733	602	618	628	633
32	787	815	812	802	699	713	736	740

Note: Response times for data points are expressed in centiseconds. Item numbers of data points correspond to numbers in Table 2.

Figure Captions

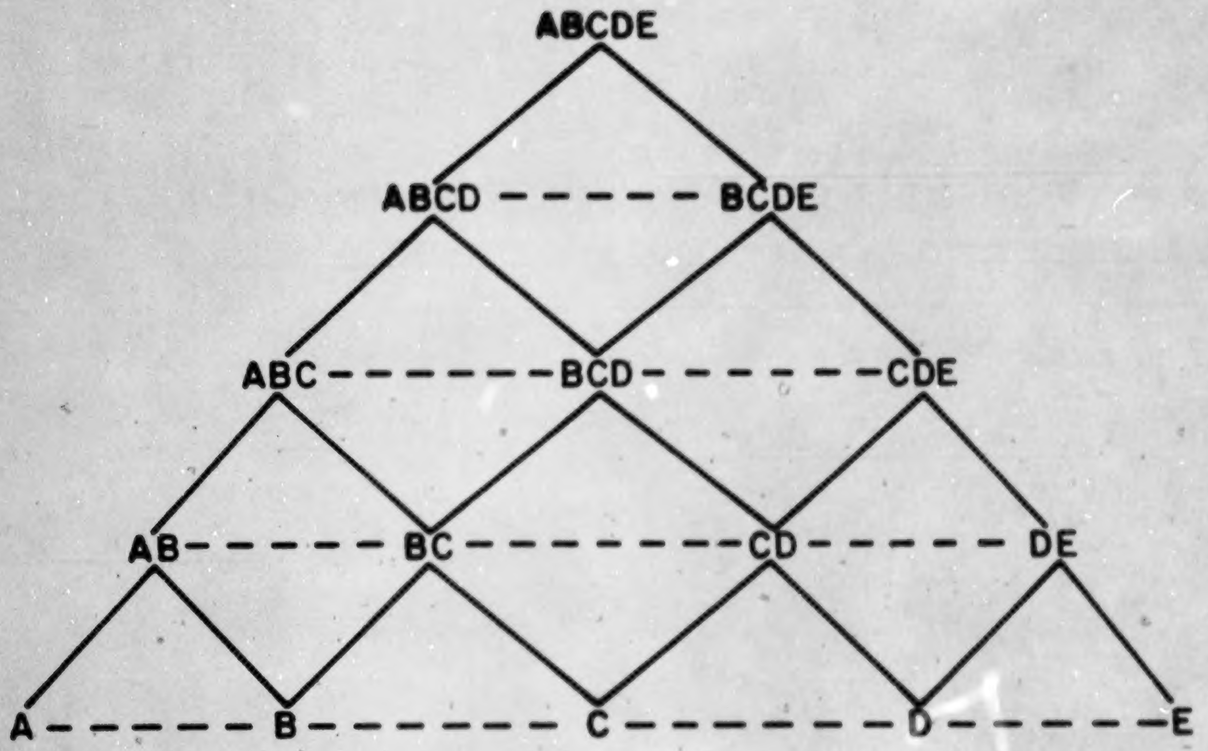
Figure 1. Hierarchical linguistic representation of information in transitive inference task.

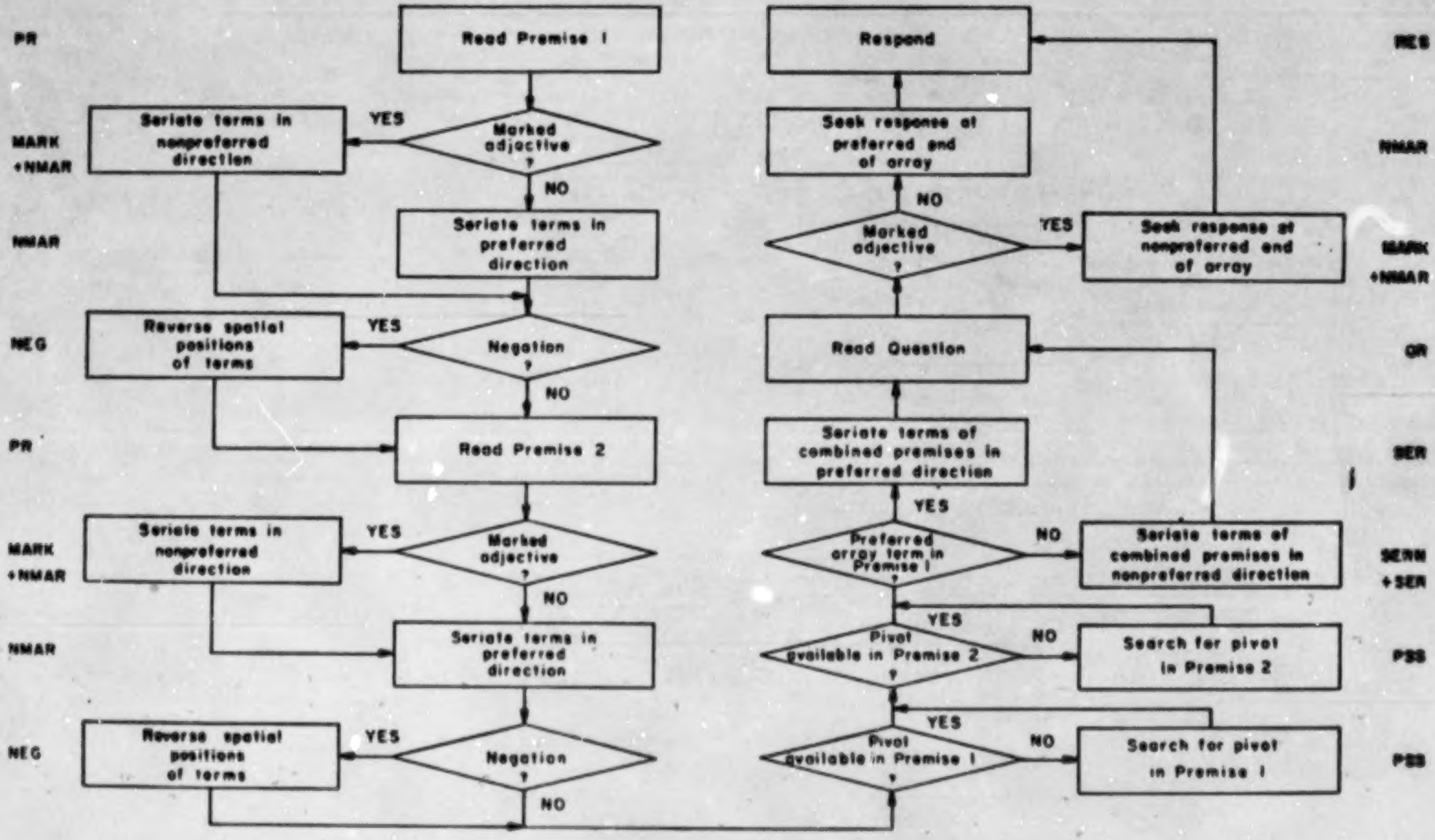
Figure 2. Spatial model.

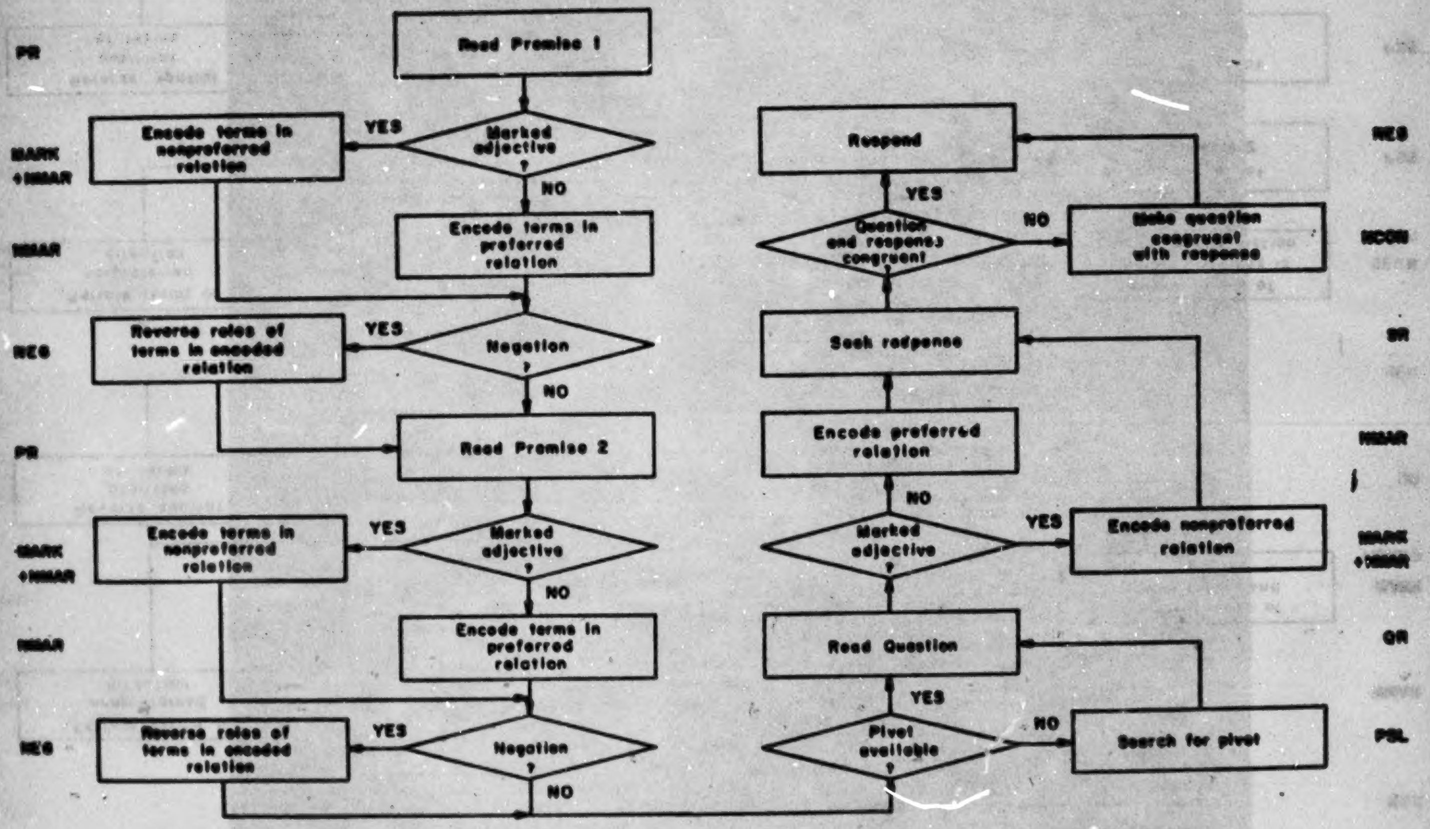
Figure 3. Linguistic model.

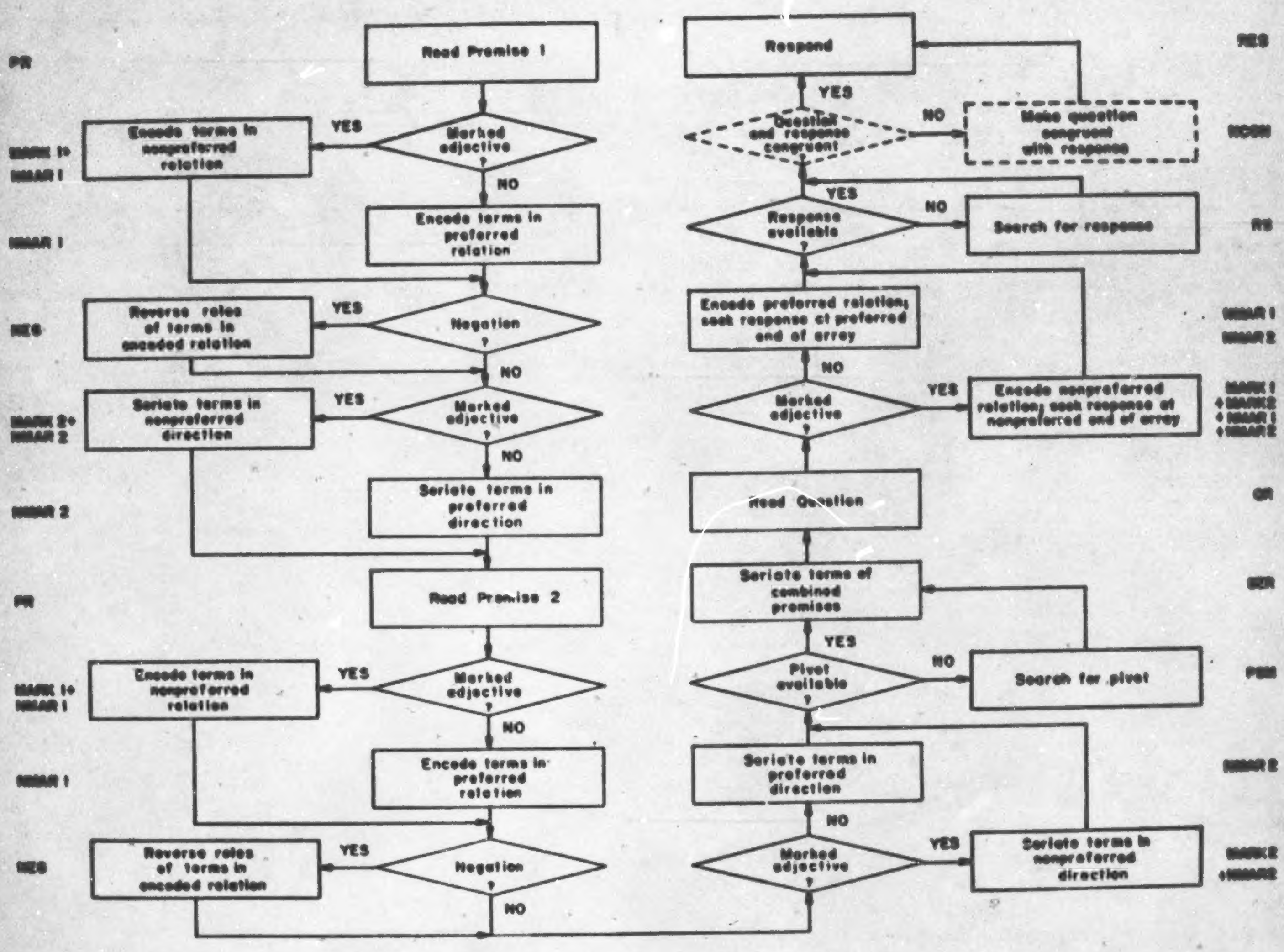
Figure 4. Mixed model.

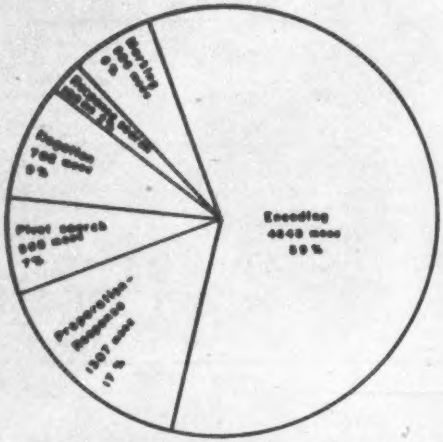
Figure 5. Amounts of time spent in each operation (or complex of operations) for a typical negative equative item in each of Experiments 1, 2, 3, and 4.


ABCDE
ABCD
BCDE
ABC
BCD
CDE
AB
BC
CD
DE
A
B
C
D
E

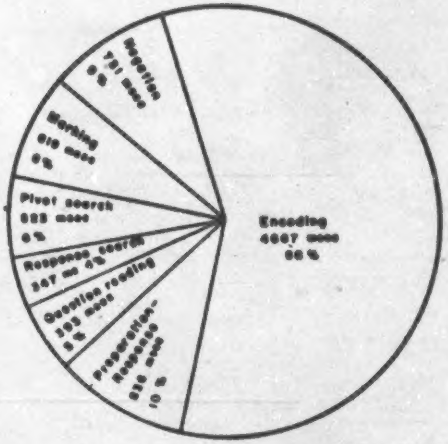

PR
Read Premise 1
MARK +NMAR
Seriate terms in nonpreferred direction
YES
Marked adjective ?
NO
NMAR
Seriate terms in preferred direction
NEG
Reverse spatial positions of terms
YES
Negation ?
NO
PR
Read Premise 2
MARK +NMAR
Seriate terms in nonpreferred direction
YES
Marked adjective ?
NO
NMAR
Seriate terms in preferred direction
NEG
Reverse spatial positions of terms
YES
Negation ?
NO
Respond
RES
Seek response at preferred end of array
NMAR
NO
Marked adjective ?
YES
Seek response at nonpreferred end of array
MARK +NMAR
Read Question
QR
Seriate terms of combined premises in preferred direction
SER
YES
Preferred array term in Premise 1 ?
NO
Seriate terms of combined premises in nonpreferred direction
SERN +SER
YES
Pivot available in Premise 2 ?
NO
Search for pivot in Premise 2
PSS
YES
Pivot available in Premise 1 ?
NO
Search for pivot in Premise 1
PSS


PR
Read Premise 1
Marked adjective ?
YES
Encode terms in nonpreferred relation
MARK +NMAR
NO
Encode terms in preferred relation
NMAR
Negation ?
YES
Reverse roles of terms in encoded relation
NEG
NO
Read Premise 2
PR
Marked adjective ?
YES
Encode terms in nonpreferred relation
MARK +NMAR
NO
Encode terms in preferred relation
NMAR
Negation ?
YES
Reverse roles of terms in encoded relation
NEG
NO
Pivot available ?
NO
Search for pivot
PSL
YES
Read Question
QR
Marked adjective ?
YES
Encode nonpreferred relation
MARK +NMAR
NO
Encode preferred relation
NMAR
Seek response
SR
Question and response congruent ?
NO
Make question congruent with response
NCON
YES
Respond
RES

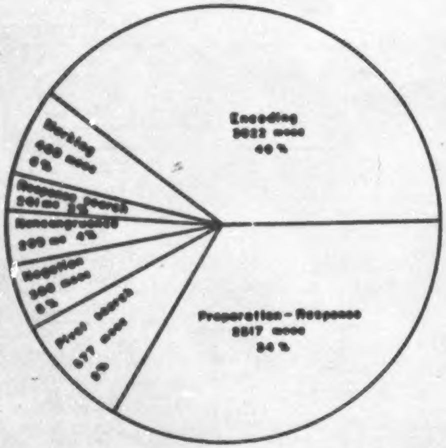

PR
Read Premise 1
MARK 1+
NMAR 1
Encode terms in nonpreferred relation
YES
Marked adjective ?
NO
NMAR 1
Encode terms in preferred relation
NEG
Reverse roles of terms in encoded relation
YES
Negation ?
NO
MARK 2+
NMAR 2
Seriate terms in nonpreferred direction
YES
Marked adjective ?
NO
NMAR 2
Seriate terms in preferred direction
PR
Read Premise 2
MARK 1+
NMAR 1
Encode terms in nonpreferred relation
YES
Marked adjective ?
NO
NMAR 1
Encode terms in preferred relation
NEG
Reverse roles of terms in encoded relation
YES
Negation ?
NO
Respond
RES
YES
Question and response congruent
NO
Make question congruent with response
NCON
YES
Response available ?
NO
Search for response
RS
Encode preferred relation; seek response at preferred end of array
NMAR 1
NMAR 2
NO
Marked adjective ?
YES
Encode nonpreferred relation; seek response at nonpreferred end of array
MARK 1
+MARK 2
+NMAR 1
+NMAR 2
Read Question
QR
Seriate terms of combined premises
SER
YES
Pivot available ?
NO
Search for pivot
PSR
Seriate terms in preferred direction
NMAR 2
NO
Marked adjective ?
YES
Seriate terms in nonpreferred direction
MARK 2
+NMAR 2



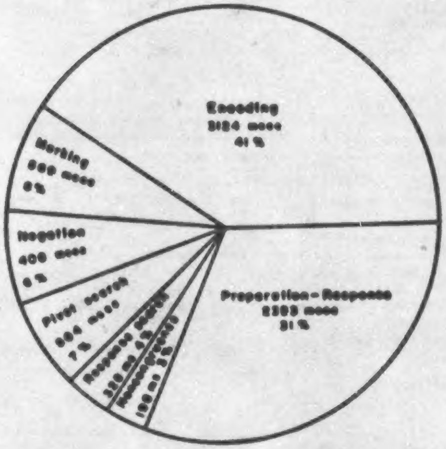
Experiment 1 7921 msec



Experiment 2 8116 msec



Experiment 3 7474 msec



Experiment 4 7564 msec

Technical Reports Presently in this Series

NR 150-412, ONR Contract N0001478C0025

#1. Sternberg, R. J. *Intelligence research at the interface between differential and cognitive psychology: Prospects and proposals*. January, 1978.

#2. Sternberg, R. J. *Isolating the components of intelligence*. January, 1978.

#3. Sternberg, R. J., Guyote, M. J., & Turner, M. E. *Deductive reasoning*. January, 1978.

#4. Sternberg, R. J. *Toward a unified componential theory of human reasoning*. April, 1978.

#5. Guyote, M. J., & Sternberg, R. J. *A transitive-chain theory of syllogistic reasoning*. April, 1978.

#6. Sternberg, R. J., & Turner, M. E. *Components of syllogistic reasoning*. April, 1978.

#7. Sternberg, R. J., Tourangeau, R., & Nigro, G. *Metaphor, induction, and social policy: The convergence of macroscopic and microscopic views*. April, 1978.

#8. Sternberg, R. J. *A proposed resolution of curious conflicts in the literature on linear syllogistic reasoning*. June, 1978.

#9. Sternberg, R. J. *The nature of mental abilities*. June, 1978.

#10. Sternberg, R. J. *Psychometrics, mathematical psychology, and cognition: Confessions of a closet psychometrician*. June, 1978.

#11. Tourangeau, R., & Sternberg, R. J. *Understanding and appreciating metaphors*. June, 1978.

#12. Sternberg, R. J. *Representation and process in transitive inference*. October, 1978.

#13. Tourangeau, R., & Sternberg, R. J. *Aptness in metaphor*. October, 1978.

#14. Sternberg, R. J. *Contrasting conceptions of intelligence and their educational implications*. November, 1978.